

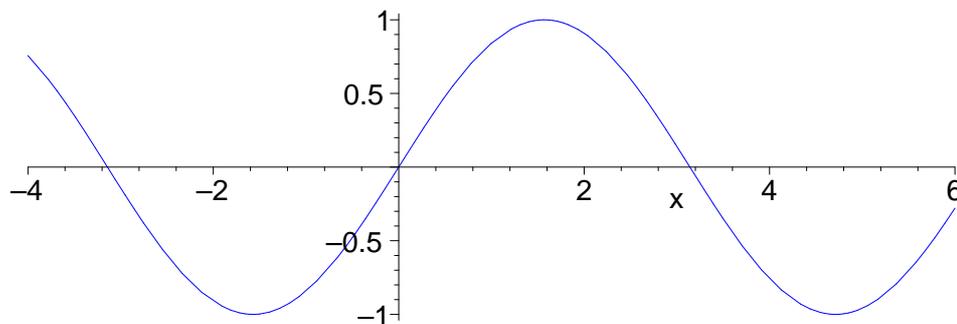
Kapitel 3

Differenzierbare Funktionen

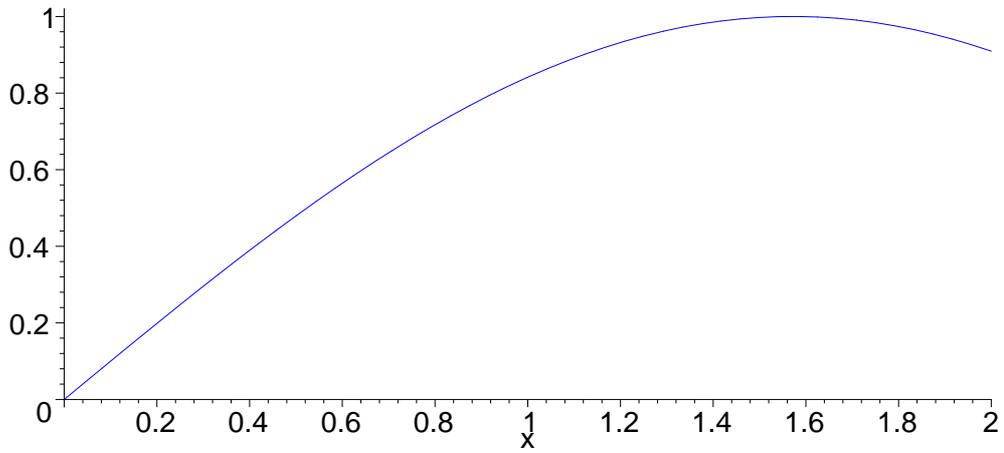
Wie zu Beginn der Vorlesung erwähnt, ist das Wort *Analysis* abgeleitet vom griechischen ἀναλύειν = auflösen, zerlegen; in diesem Kapitel werden wir nun endlich Funktionen untersuchen, indem wir sie in ihre kleinsten Bestandteile zerlegen, d.h. indem wir ihr Verhalten in der unmittelbaren Umgebung eines Punktes untersuchen. In diesem Kapitel geht es um Funktionen, die sich dort „fast“ wie lineare Funktionen verhalten, die sogenannten differenzierbaren Funktionen.

§ 1: Differenzierbarkeit

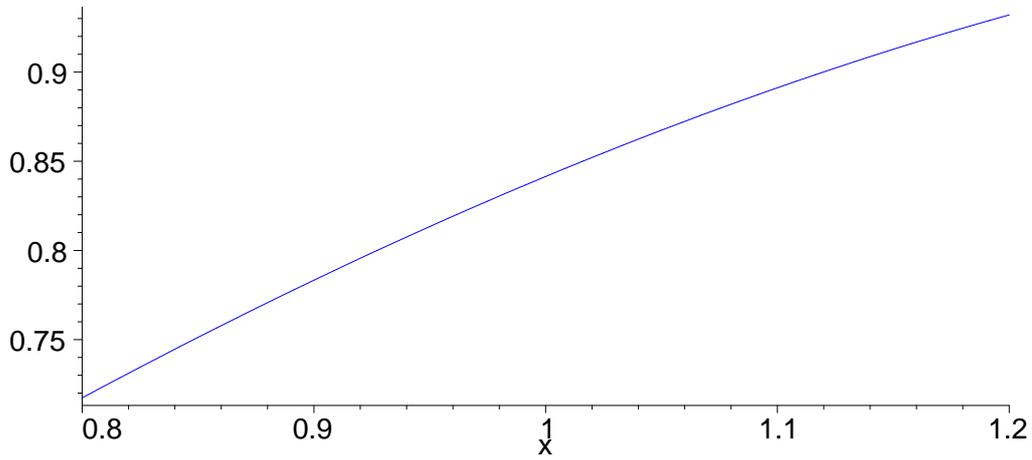
Um zu verstehen, was „fast“ linear bedeuten soll, betrachten wir die Graphen dreier Funktionen und vergrößern sie in der Umgebung eines Punktes. Das erste Beispiel ist die Funktion $f(x) = \sin x$ in der Umgebung des Punktes $x = 1$. (Wer diese Funktion nicht aus der Schule kennt, kann sie einfach als irgendeine Funktion ansehen.) Wie die drei nachfolgenden Bilder zeigen, nähert sich das Bild beim Hereinzoomen immer mehr einer Geraden; beim letzten Bild ist zumindest visuell kein Unterschied zu einer echten Gerade zu erkennen. So etwas ist der Prototyp einer differenzierbaren Funktion.



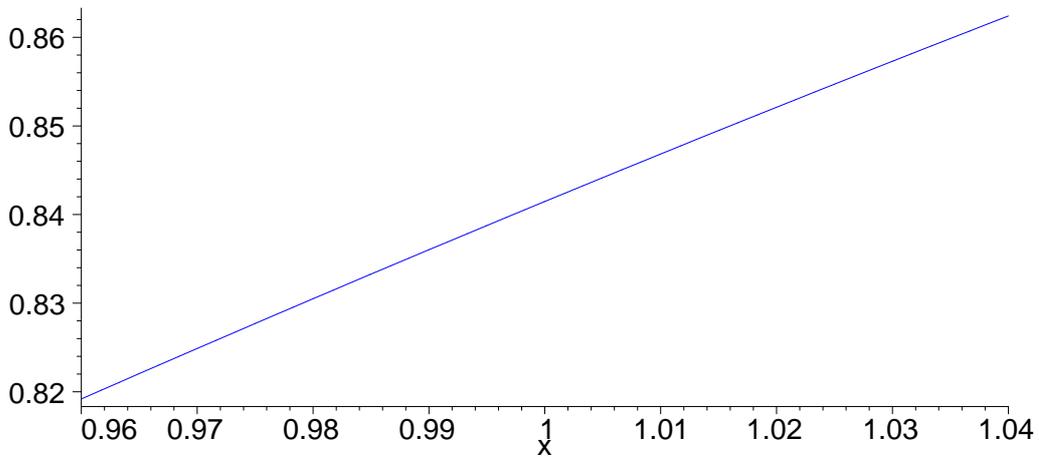
Graph der Funktion $y = \sin x$



Vergrößerung um den Faktor fünf um $(1, \sin(1))$

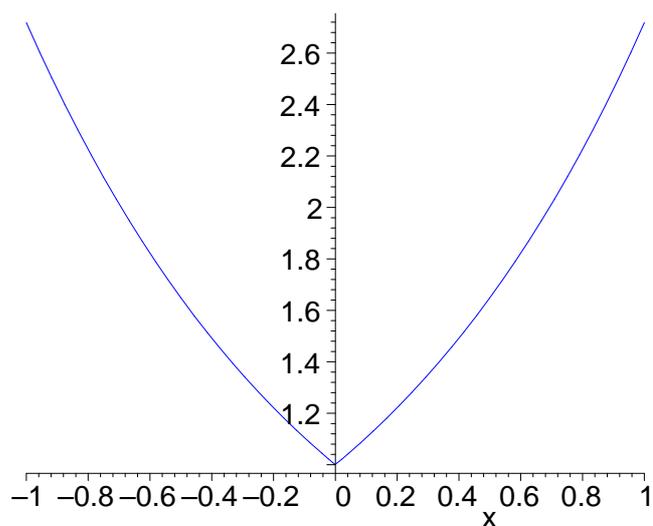


Nochmalige Vergrößerung um den Faktor fünf um $(1, \sin(1))$

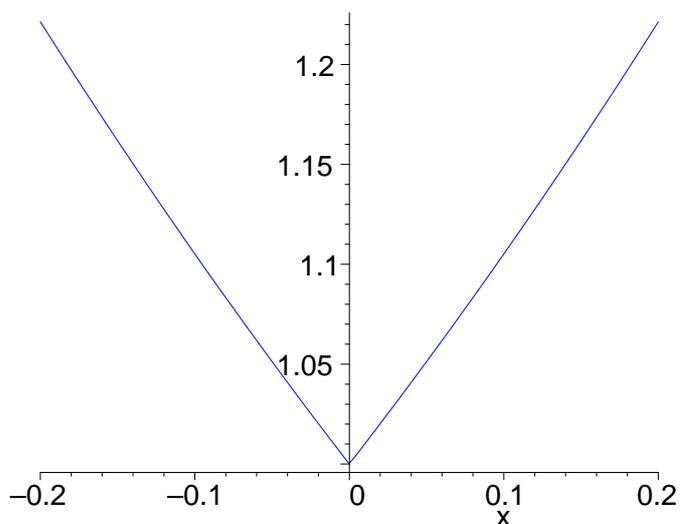


Nach noch einer Vergrößerung sieht die Funktion praktisch linear aus

Als zweites Beispiel betrachten wir die Funktion $f(x) = e^{|x|}$. Beim Vergrößern werden zwar die beiden „Zweige“ des Graphen zunehmend linearer, der „Knick“ bleibt aber. Dies ist ein Beispiel einer Funktion, die in *einem* speziellen Punkt nicht differenzierbar ist.



Graph der Funktion $y = e^{|x|}$

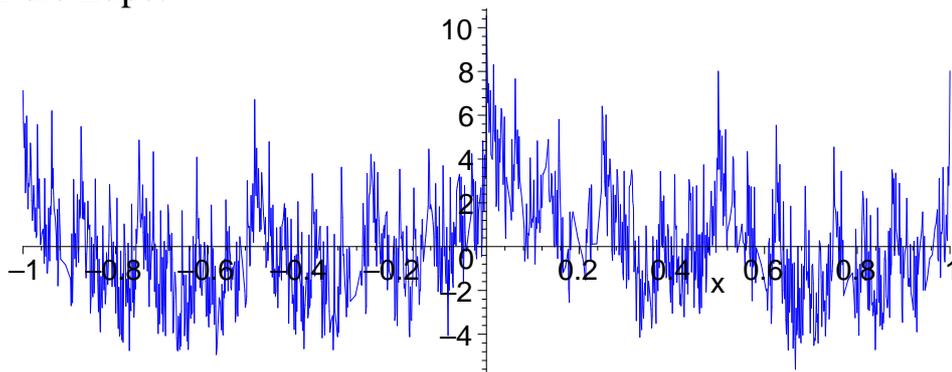


Vergrößerung um den Faktor fünf

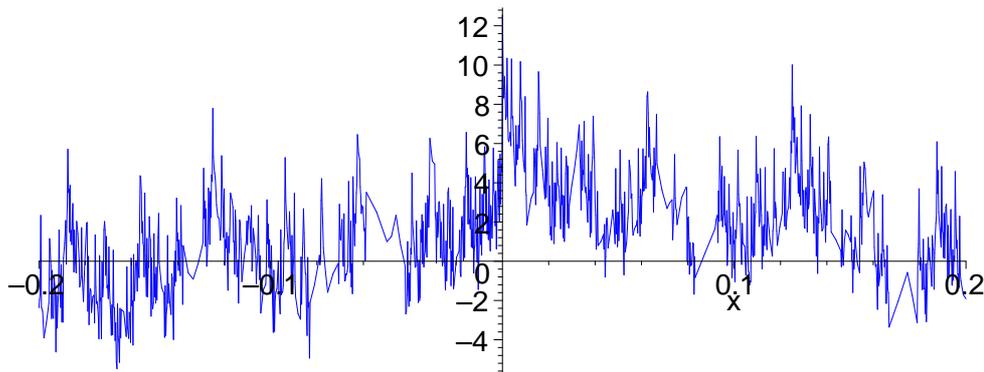
Als letztes Beispiel nehmen wir noch den Graphen der Funktion

$$f(x) = \sum_{k=0}^{\infty} \cos([2^k \pi x])$$

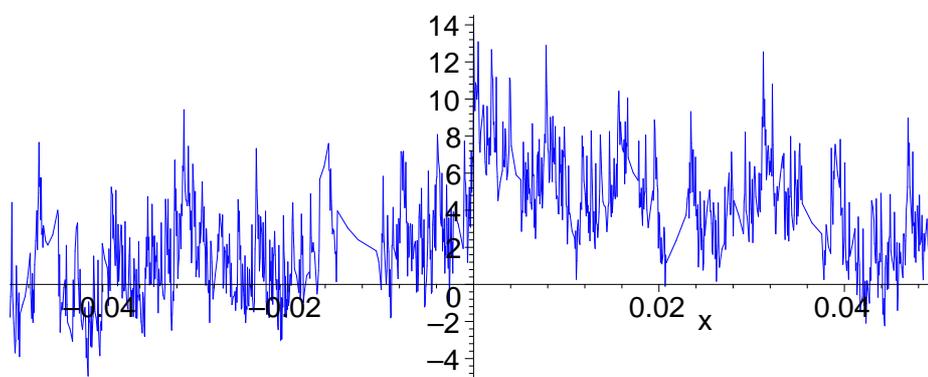
unter die Lupe:



Graph von f



Vergrößerung um den Faktor fünf



Vergrößerung um den Faktor zwanzig

Die drei Bilder sind zwar durchaus verschieden, aber im wesentlichen sehen wir in jedem Vergrößerungsmaßstab das gleiche Chaos. So etwas ist ein Beispiel einer *nirgends* differenzierbaren Funktion; die hier beobachtete Skaleninvarianz ist typisch für alle Arten von Rauschen und auch für die Entwicklung vieler Börsenkurse, wo das Bild der im Minutenrhythmus aufgezeichneten Kursentwicklung innerhalb einer Stunde ohne Beschriftung oft nicht vom Diagramm der Kursentwicklung des letzten Jahres unterscheidbar ist.

Wir konzentrieren uns in diesem Kapitel auf Funktionen des ersten Typs und müssen diese daher exakt definieren. Die Grundidee ist, daß sich die Funktion in der Umgebung eines jeden Punkts „fast“ wie eine lineare Funktion verhalten soll; für kleine Werte von h soll also gelten

$$f(x_0 + h) \approx f(x_0) + ah \quad \text{mit einem geeigneten } a \in \mathbb{R}.$$

Diese Gleichung kann für eine nichtlineare Funktion natürlich nicht exakt gelten; um eine Gleichheit zu bekommen, müssen wir noch einen Fehlerterm einführen. Dieser muß für $h = 0$ verschwinden, denn hier haben wir ja stets Gleichheit, und für kleine Werte von h soll er kleiner sein als der lineare Term ah . Um dies zu erreichen, verlangen wir, daß er sich in der Form $h \cdot \tilde{f}(h)$ schreiben läßt, wobei $\tilde{f}(h)$ eine stetige Funktion ist, die an der Stelle $h = 0$ verschwindet. Für eine exakte Definition müssen wir noch darauf achten, daß die Funktion im Punkt $x_0 + h$ überhaupt definiert ist und daß es genügend viele kleine $h \neq 0$ gibt, für die $f(x_0 + h)$ definiert ist. Letzteres Problem lösen wir dadurch, daß wir verlangen, daß es in jeder Umgebung von x_0 unendlich viele Punkte von D gibt; formal:

Definition: Ein Punkt x_0 aus der Teilmenge $D \subseteq \mathbb{R}$ heißt *Häufungspunkt* von D , wenn es für jedes $\varepsilon > 0$ unendlich viele Punkte $x \in D$ gibt mit $|x - x_0| < \varepsilon$.

In einem Intervall D , egal ob offen oder abgeschlossen, das nicht nur aus einem Punkt besteht, ist jeder Punkt ein Häufungspunkt; das ist die Situation, mit der wir es typischerweise zu tun haben werden. In $D = \mathbb{Z}$ dagegen ist *kein* Punkt Häufungspunkt, denn für $\varepsilon < 1$ gibt es kein $x \neq x_0$ in \mathbb{Z} mit $|x - x_0| < \varepsilon$. Aus diesem Grund redet auch niemand von differenzierbaren Funktionen auf \mathbb{Z} .

In voller Allgemeinheit sieht unsere Definition der Differenzierbarkeit nun folgendermaßen aus:

Definition: $f: D \rightarrow \mathbb{R}$ sei definiert auf der Teilmenge $D \subseteq \mathbb{R}$ und x_0 sei ein Häufungspunkt von D . f heißt *differenzierbar* in x_0 , wenn es eine reelle Zahl a gibt sowie ein $\varepsilon > 0$ und für

$$D_{x_0, \varepsilon} \stackrel{\text{def}}{=} \{h \in \mathbb{R} \mid x_0 + h \in D \text{ und } |h| < \varepsilon\}$$

eine stetige Funktion $\tilde{f}: D_{x_0, \varepsilon} \rightarrow \mathbb{R}$ mit $\tilde{f}(0) = 0$, so daß

$$f(x_0 + h) = f(x_0) + ah + h\tilde{f}(h) \quad \text{für alle } h \in D_{x_0, \varepsilon}.$$

Die Zahl a wird als die Ableitung $a = f'(x_0)$ von f im Punkt x_0 bezeichnet. f heißt differenzierbar auf D , wenn f in jedem Punkte $x_0 \in D$ differenzierbar ist.

Im (häufigen) Fall, daß es ein Intervall $(x_0 - \delta, x_0 + \delta)$ gibt, das ganz in D liegt, können wir das erheblich einfacher formulieren: Dann muß es einfach ein $a \in \mathbb{R}$ und ein $\varepsilon > 0$ geben, so daß $f(x_0 + h) = f(x_0) + ah + h\tilde{f}(h)$ ist für alle h mit $|h| < \varepsilon$. Falls $D = [x_1, x_2]$ ein abgeschlossenes Intervall ist, bedeutet Differenzierbarkeit im Anfangspunkt x_1 , daß es a und $\varepsilon > 0$ gibt, so daß $f(x_1 + h) = f(x_1) + ah + h\tilde{f}(h)$ für alle $h \in [0, \varepsilon)$; für die Differenzierbarkeit in x_2 muß es entsprechend ein $b \in \mathbb{R}$ geben, so daß $f(x_2 + h) = f(x_2) + bh + h\tilde{f}(h)$ ist für alle $h \in (-\varepsilon, 0]$.

Geometrisch gesprochen ist eine Funktion f in einem Punkt x_0 differenzierbar, wenn ihr Graph dort eine Tangente hat, nämlich die Gerade $y = f(x_0) + a(x - x_0)$. Dies ist die einzige Gerade durch den Punkt $(x_0, f(x_0))$, für die der Abstand zwischen Kurven- und Geradenpunkt für $x \rightarrow x_0$ schneller gegen Null geht als eine lineare Funktion in $x - x_0$.

In der Schule wird die Ableitung einer Funktion f meist definiert als der Grenzwert

$$f'(x_0) = \lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0)}{h}.$$

Um den Zusammenhang damit zu sehen, müssen wir uns zunächst überlegen, was der rechts stehende Grenzwert bedeuten soll: Bislang kennen wir nur Limites von Folgen.

Definition: $f: D \rightarrow \mathbb{R}$ sei eine Funktion und x_0 ein Häufungspunkt der Menge D . Dann ist

$$\lim_{x \rightarrow x_0} f(x) = y_0$$

genau dann, wenn es zu jedem $\varepsilon > 0$ ein $\delta > 0$ gibt, so daß gilt: $|f(x) - y_0| < \varepsilon$ für alle $x \in D$ mit $|x - x_0| < \delta$.

Diese Definition erinnert sehr an die der Stetigkeit, und in der Tat überlegt man sich leicht, daß f genau dann stetig in x_0 ist, wenn $\lim_{x \rightarrow x_0} f(x) = f(x_0)$ ist.

Für eine differenzierbare Funktion, wie sie hier definiert wurde, ist

$$\frac{f(x_0 + h) - f(x_0)}{h} = \frac{ah + h\tilde{f}(h)}{h} = a + \tilde{f}(h),$$

was für $h \rightarrow 0$ in der Tat gegen die Ableitung $a = f'(x_0)$ konvergiert. Umgekehrt ist

$$f(x_0 + h) = f(x_0) + h \cdot \frac{f(x_0 + h) - f(x_0)}{h};$$

falls der Limes des Differenzenquotienten gegen eine Zahl a konvergiert, können wir dies auch schreiben als

$$f(x_0 + h) = f(x_0) + ah + h \left(\frac{f(x_0 + h) - f(x_0)}{h} - a \right),$$

wobei die Funktion

$$\tilde{f}(h) = \frac{f(x_0 + h) - f(x_0)}{h} - a$$

stetig ist und für $h = 0$ verschwindet. Somit sind die beiden Definitionen äquivalent.

Betrachten wir einige Beispiele!

Für die Funktion $f(x) = x^n$ ist

$$\begin{aligned} f(x_0 + h) &= (h + x_0)^n = \sum_{k=0}^n \binom{n}{k} h^k x_0^{n-k} \\ &= x_0^n + \binom{n}{1} h x_0^{n-1} + h \sum_{k=2}^n \binom{n}{k} h^{k-1} x_0^{n-k}, \end{aligned}$$

also ist

$$f(x_0+h) = f(x_0) + nx_0^{n-1} \cdot h + h\tilde{f}(h) \quad \text{mit} \quad \tilde{f}(h) = \sum_{k=2}^n \binom{n}{k} h^{k-1} x_0^{n-k}.$$

Da in jedem Summanden der hinteren Summe mindestens ein h steckt, ist $\tilde{f}(0) = 0$, und als Polynom in h ist \tilde{f} natürlich stetig. Somit ist f differenzierbar im Punkt x_0 und hat dort die Ableitung nx_0^{n-1} . Der Punkt x_0 war beliebig gewählt, also ist die Funktion differenzierbar auf ganz \mathbb{R} und $f'(x) = nx^{n-1}$.

Als zweites Beispiel betrachten wir die Exponentialfunktion. Hier ist

$$e^{x+h} = e^x \cdot e^h = e^x + (e^h - 1)e^x.$$

Wegen der Bedingungen aus der Definition der Exponentialfunktion ist

$$1 + h \leq e^h \leq \frac{1}{1-h},$$

also

$$\begin{aligned} h \leq e^h - 1 &\leq \frac{1}{1-h} - 1 = \frac{1 - (1-h)}{1-h} = \frac{h}{1-h} \\ &= \frac{h(1-h) + h^2}{1-h} = h + \frac{h^2}{1-h}. \end{aligned}$$

Mit

$$\tilde{f}(h) = \frac{e^{x+h} - e^x}{h} - e^x = e^x \left(\frac{e^h - 1}{h} - 1 \right)$$

ist daher

$$e^{x+h} = e^x + he^x + h\tilde{f}(h) \quad \text{mit} \quad 0 \leq \tilde{f}(h) \leq \frac{he^x}{1-h}.$$

Insbesondere ist $\tilde{f}(0) = 0$, und $\tilde{f}(h)$ ist stetig an der Stelle $h = 0$, denn für jede Nullfolge $(x_n)_{n \in \mathbb{N}}$ muß auch die Folge der $\tilde{f}(x_n)$ wegen dieser Ungleichungen gegen Null konvergieren.

Damit ist auch die Exponentialfunktion differenzierbar auf ganz \mathbb{R} , und sie ist ihre eigene Ableitung.

Um möglichst schnell möglichst viele weitere Beispiele zu bekommen, brauchen wir Rechenregeln für Ableitungen.

Die erste folgt fast sofort aus der Definition:

Summenformel: Sind $f, g: D \rightarrow \mathbb{R}$ differenzierbare Funktionen, so ist auch die Summenfunktion $f + g$ differenzierbar und

$$(f + g)'(x) = f'(x) + g'(x).$$

Beweis: Da f und g differenzierbar sind, haben wir für jedes $x_0 \in \mathbb{R}$ zwei reelle Zahlen $a = f'(x_0)$ und $b = g'(x_0)$ sowie zwei Funktionen $\tilde{f}(h), \tilde{g}(h)$, so daß gilt:

$$f(x_0 + h) = f(x_0) + ah + h\tilde{f}(h) \quad \text{und} \quad g(x_0 + h) = g(x_0) + bh + h\tilde{g}(h).$$

Somit ist

$$(f+g)(x_0+h) = f(x_0+h)+g(x_0+h) = (f+g)(x_0)+(a+b)h+h \cdot (\tilde{f}(h)+\tilde{g}(h)).$$

Da $\tilde{f}(h) + \tilde{g}(h)$ als Summe zweier stetiger Funktionen stetig ist und an der Stelle $h = 0$ verschwindet, ist $f + g$ in x_0 differenzierbar mit Ableitung $a + b = f'(x_0) + g'(x_0)$, wie behauptet. ■

Genauso einfach folgt

Lemma: Ist $f: D \rightarrow \mathbb{R}$ differenzierbar und $c \in \mathbb{R}$ eine Konstante, so ist auch die Funktion cf , die jedes $x \in D$ auf $cf(x)$ abbildet, differenzierbar und hat die Ableitung $cf'(x)$.

Zum *Beweis* müssen wir nur in der Definition der Differenzierbarkeit die Gleichung mit c multiplizieren. ■

Die beiden vorigen Resultate können wir zusammenfassen, indem wir sagen, daß Differenzieren eine *lineare* Operation ist, d.h. für zwei differenzierbare Funktionen f, g und zwei reelle Zahlen a, b ist

$$(af + bg)'(x) = af'(x) + bg'(x),$$

und entsprechend natürlich auch für mehr als zwei Summanden.

Mit dieser Regel können wir insbesondere Polynome ableiten:

Lemma: Jedes Polynom $f(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0$ definiert eine differenzierbare Funktion $f: \mathbb{R} \rightarrow \mathbb{R}$ mit Ableitung

$$f'(x) = na_n x^{n-1} + (n-1)a_{n-1} x^{n-2} + \dots + 2a_2 x + a_1.$$

Beweis: Wir wissen, daß x^k auf ganz \mathbb{R} differenzierbar ist mit Ableitung kx^{k-1} ; wegen der Linearität des Ableitens folgt daraus die Behauptung. ■

Um auch Funktionen wie $x \cdot e^x$ differenzieren zu können, müssen wir nach den Summen auch Produkte differenzierbarer Funktionen untersuchen; hier gilt die

Leibnizsche Produktregel: Sind $f, g: D \rightarrow \mathbb{R}$ differenzierbare Funktionen, so ist auch die Produktfunktion $f \cdot g$ differenzierbar und

$$(fg)'(x) = f'(x)g(x) + f(x)g'(x).$$

Der *Beweis* folgt auch hier sofort durch Einsetzen: Ist für einen Punkt $x_0 \in D$

$$f(x_0 + h) = f(x_0) + hf'(x_0) + h\tilde{f}(h)$$

und

$$g(x_0 + h) = g(x_0) + hg'(x_0) + h\tilde{g}(h),$$

so ist

$$\begin{aligned} (fg)(x_0 + h) &= (f(x_0) + hf'(x_0) + h\tilde{f}(h))(g(x_0) + hg'(x_0) + h\tilde{g}(h)) \\ &= (fg)(x_0) + h(f'(x_0)g(x_0) + f(x_0)g'(x_0)) \\ &\quad + h\left(\tilde{f}(h)g(x_0 + h) + \tilde{g}(h)(f(x_0) + hf'(x_0))\right). \end{aligned}$$

Die Funktion in der letzten Klammer ist stetig, da sie aus stetigen Funktionen zusammengesetzt ist, und sie verschwindet für $h = 0$, da $\tilde{f}(h) = \tilde{g}(h) = 0$ ist. Also ist fg in x_0 differenzierbar mit der behaupteten Ableitung, und dies gilt für jeden Punkt $x_0 \in D$. ■

Die nächste wichtige Differentiationsregel behandelt die Hintereinanderausführung von Funktionen; hier gilt

Kettenregel: Sind $f: D \rightarrow \mathbb{R}$ und $g: E \rightarrow \mathbb{R}$ differenzierbare Funktionen und ist $g(x) \in D$ für alle $x \in E$, so ist auch die Hintereinanderausführung $f \circ g: E \rightarrow \mathbb{R}$ differenzierbar, und

$$(f \circ g)'(x) = f'(g(x)) \cdot g'(x).$$

Beweis: Dieses Mal drücken wir die Differenzierbarkeit von f und g so aus, daß

$$g(x_0 + h) = g(x_0) + hg'(x_0) + h\tilde{g}(h)$$

und

$$f(u_0 + k) = f(u_0) + kf'(u_0) + k\tilde{f}(k)$$

ist. Dann ist für jedes $x_0 \in E$

$$(f \circ g)(x_0 + h) = f(g(x_0 + h)) = f(g(x_0) + hg'(x_0) + h\tilde{g}(h)).$$

Setzen wir $u_0 = g(x_0)$ und $k = hg'(x_0) + h\tilde{g}(h)$ in die Formel für $f(u_0 + k)$ ein, können wir dies weiter ausrechnen als

$$f(g(x_0 + h)) = f(g(x_0)) + k \cdot f'(g(x_0)) + k\tilde{f}(k).$$

Da mit h auch k gegen Null geht, folgt die Behauptung. ■

Damit können wir nun auch Funktionen wie $f(x) = e^{-x^2}$ differenzieren; die Ableitung ist $f'(x) = -2xe^{-x^2}$. Weitere Beispiele für Anwendungen der gerade bewiesenen Regeln werden in den Übungen sowie auch im weiteren Verlauf der Vorlesung auftauchen.

Auch Quotienten sind kein Problem; nur müssen wir hier natürlich voraussetzen, daß der Nenner nirgends verschwindet, was wir bei konkreten Anwendungen dadurch erreichen können, daß wir den Definitionsbereich D entsprechend verkleinern.

Quotientenregel: Sind $f, g: D \rightarrow \mathbb{R}$ differenzierbare Funktionen und ist $g(x) \neq 0$ für alle $x \in D$, so ist auch der Quotient f/g differenzierbar und

$$\left(\frac{f}{g}\right)'(x) = \frac{f'(x)g(x) - f(x)g'(x)}{g(x)^2}.$$

Beweis: Da wir die Produktregel bereits bewiesen haben, genügt es, diese Regel im Spezialfall $f \equiv 1$ zu beweisen; hier besagt sie, daß

$$\left(\frac{1}{g}\right)'(x) = -\frac{g'(x)}{g(x)^2}$$

ist. Wenn wir dies wissen, folgt aus der Produktregel, daß

$$\begin{aligned} \left(\frac{f}{g}\right)'(x) &= \left(f \cdot \frac{1}{g}\right)'(x) = f'(x) \left(\frac{1}{g}\right)(x) + f(x) \left(\frac{1}{g}\right)'(x) \\ &= \frac{f'(x)}{g(x)} - \frac{f(x)g'(x)}{g(x)^2} = \frac{f'(x)g(x) - f(x)g'(x)}{g(x)^2} \end{aligned}$$

ist. Bleibt also die Ableitung des Kehrwerts.

Betrachten wir als erstes die Funktion $f(x) = 1/x$ für ein $x_0 \neq 0$. Nach den Regeln der Bruchrechnung ist

$$\frac{1}{x+h} - \frac{1}{x} = \frac{-h}{x(x+h)}, \quad \text{also} \quad \frac{1}{x+h} = \frac{1}{x} - \frac{h}{x(x+h)} = \frac{1}{x} - \frac{h}{x} \frac{1}{x+h}.$$

Wenden wir für den letzten Faktor die Formel rekursiv an, erhalten wir die neue Formel

$$\frac{1}{x+h} = \frac{1}{x} - \frac{h}{x} \left(\frac{1}{x} - \frac{h}{x(x+h)} \right) = \frac{1}{x} - \frac{h}{x^2} + \frac{h^2}{x^2(x+h)}.$$

Der letzte Summand hat die Form $h\tilde{f}(h)$ mit einer stetigen Funktion \tilde{f} , die für $h = 0$ verschwindet, also ist $f'(x) = -1/x^2$ die Ableitung von f .

Für die Nennerfunktion g aus der Behauptung können wir $1/g$ auffassen als Hintereinanderausführung von g und der Funktion $x \mapsto 1/x$; nach der Kettenregel ist daher

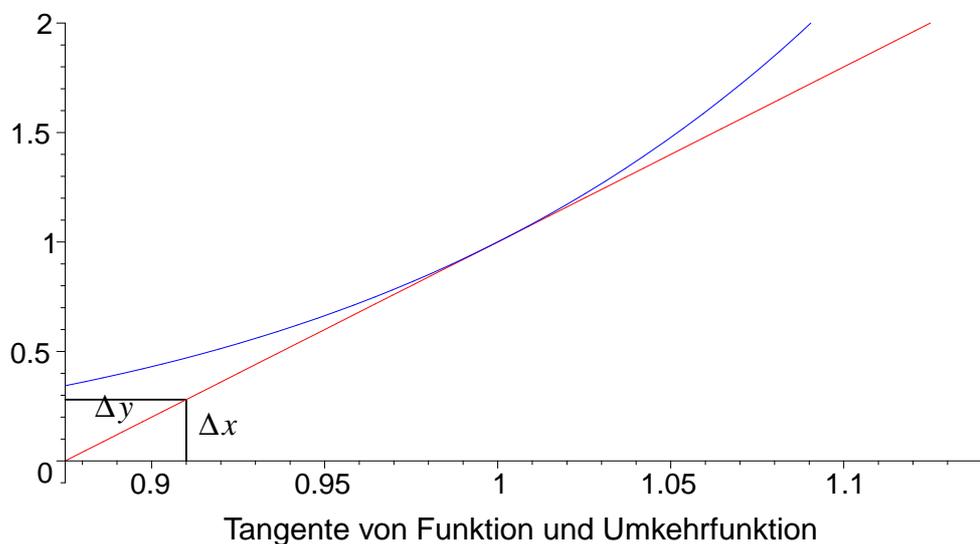
$$\left(\frac{1}{g}\right)'(x) = -\frac{1}{g(x)^2} \cdot g'(x) = -\frac{g'(x)}{g(x)^2},$$

wie behauptet. ■

Was uns nun noch fehlt von den bisher bekannten Funktionen sind zum Beispiel der (natürliche) Logarithmus und die Wurzelfunktion. Beides sind Umkehrfunktionen von Funktionen, deren Ableitung wir bereits kennen; wenn wir also wissen, wie man Umkehrfunktionen differenziert, können wir auch diese Funktionen ableiten.

Geometrisch betrachtet ist die Situation ganz einfach: Wenn eine Funktion $f: D \rightarrow \mathbb{R}$ eine Umkehrfunktion hat (was natürlich nur dann der Fall ist, wenn f injektiv ist), entsteht deren Graph aus dem der Funktion f einfach dadurch, daß x - und y -Achse vertauscht werden. Die

Ableitung einer Funktion in einem gegebenen Punkt ist die Steigung der Tangente des Graphen; wenn man x - und y -Achse vertauscht, wird die Steigung nicht mehr in Bezug auf die x -Achse, sondern in Bezug auf die y -Achse gemessen. In Bezug auf die x -Achse ist die Steigung einer Geraden der Quotient aus einer (jeder von Null verschiedener) Differenz von y -Werten und der zugehörigen Differenz von x -Werten; in Bezug auf die y -Achse ist es umgekehrt. Die Ableitung der Umkehrfunktion sollte also einfach der Kehrwert der Ableitung der Funktion im gegebenen Punkt sein.



Bevor wir dies formal beweisen können, müssen wir zunächst sicherstellen, daß es überhaupt eine Umkehrfunktion gibt, daß also die Ausgangsfunktion injektiv ist. Auf einem Intervall ist dies, wie wir im vorigen Kapitel gesehen haben, äquivalent zur strengen Monotonie. Für streng monotone Funktionen gilt

Ableitung der Umkehrfunktion: Ist $f: D \rightarrow \mathbb{R}$ streng monoton und differenzierbar mit $f'(x) \neq 0$ für alle $x \in D$, so ist auch die Umkehrfunktion $g: E \rightarrow D$ differenzierbar. Im Punkt $y = f(x) \in E$ ist

$$g'(y) = \frac{1}{f'(x)}.$$

Beweis: Wir betrachten die Funktion f in der Umgebung des Punktes

(x_0, y_0) mit $y_0 = f(x_0)$. Wegen der Differenzierbarkeit von f ist

$$f(x_0 + h) = f(x_0) + hf'(x_0) + h\tilde{f}(h) = y_0 + hf'(x_0) + h\tilde{f}(h)$$

mit einer stetigen Funktion \tilde{f} von h , die im Nullpunkt verschwindet. Setzen wir $x = x_0 + h$ und $y = f(x)$, können wir dies auch schreiben als

$$y = y_0 + hf'(x_0) + h\tilde{f}(h).$$

Somit ist

$$h = \frac{y - y_0 - h\tilde{f}(h)}{f'(x_0)},$$

also ist

$$g(y) = x_0 + h = g(y_0) + \frac{y - y_0 - h\tilde{f}(h)}{f'(x_0)} = g(y_0) + \frac{y - y_0}{f'(x_0)} - \frac{h\tilde{f}(h)}{f'(x_0)}.$$

Setzen wir noch $k = y - y_0$ bekommen wir die Gleichung

$$g(y_0 + k) = g(y_0) + \frac{k}{f'(x_0)} - \frac{h\tilde{f}(h)}{f'(x_0)}.$$

Da wegen der Stetigkeit von f mit h auch k gegen Null geht, ist der letzte Term in der Form $k\tilde{g}(k)$ darstellbar mit einer stetigen Funktion \tilde{g} , also ist g differenzierbar in y_0 mit Ableitung $1/f'(x_0)$. ■

Als erste Anwendung können wir die Logarithmusfunktion ableiten. Sie ist die Umkehrfunktion der Exponentialfunktion; im Punkt $y = e^x$ hat $\log y$ daher die Ableitung $\frac{1}{e^x} = \frac{1}{y}$.

Auch die Wurzelfunktion können wir so behandeln, müssen allerdings den Punkt $x = 0$ ausschließen, da die Ableitung von $f(x) = x^2$ dort verschwindet. Wir betrachten sie daher nur als Umkehrfunktion von

$$f: \begin{cases} \mathbb{R}_{>0} \rightarrow \mathbb{R}_{>0} \\ x \mapsto x^2 \end{cases}$$

und sehen, daß \sqrt{y} im Punkt $y > 0$ die Ableitung

$$\frac{1}{2x} = \frac{1}{2\sqrt{y}}$$

hat. Die dritte Wurzel ist entsprechend die Umkehrfunktion der dritten Potenz, wobei wir diese auf ganz \mathbb{R} betrachten können. Wenn wir den

gerade bewiesenen Satz anwenden wollen, müssen wir allerdings auch hier den Nullpunkt ausschließen, denn dort verschwindet die Ableitung von x^3 . Für einen Punkt $y \in \mathbb{R} \setminus \{0\}$ ist dann die Ableitung von $\sqrt[3]{y}$ gleich

$$\frac{1}{3x^2} = \frac{1}{3(\sqrt[3]{y})^2} = \frac{1}{3}y^{-2/3}.$$

Für $y = 0$ würden wir hier durch Null dividieren; dort ist die Funktion also nicht differenzierbar. Entsprechend folgt, daß für jede ungerade natürliche Zahl n die n -te Wurzel auf $\mathbb{R} \setminus \{0\}$ differenzierbar ist und für jede gerade natürliche Zahl auf $\mathbb{R}_{>0}$; die Ableitung ist jeweils $\frac{1}{n}y^{(1-n)/n}$.

Tatsächlich können wir für positive Argumente sogar beliebige Potenzen ableiten: Für reelles $r \neq 0$ und $x > 0$ ist $f(x) = x^r = e^{r \log x}$, und das hat nach der Kettenregel die Ableitung

$$f'(x) = e^{r \log x} \cdot \frac{r}{x} = \frac{rx^r}{x} = rx^{r-1},$$

genau wie wir es bereits für natürliche Zahlen r und gerade eben für $r = \frac{1}{n}$ gezeigt haben.

Auch die Funktion $f(x) = a^x$ für eine reelle Zahl $a > 0$ läßt sich so nach der Kettenregel ableiten: Hier ist $f(x) = e^{x \log a}$ und

$$f'(x) = e^{x \log a} \cdot \log a = a^x \log a.$$

§2: Eigenschaften und Anwendungen differenzierbarer Funktionen

Nachdem wir nun Regeln kennen, mit denen wir viele gängige Funktionen ableiten können, sollten wir uns langsam überlegen, wozu differenzierbare Funktionen und ihre Ableitungen eigentlich gut sind. Wir hatten uns im letzten Kapitel ausführlich mit stetigen Funktionen beschäftigt; als erstes stellt sich daher die Frage, wie sich diese beiden Funktionsklassen zueinander verhalten. Die Antwort ist auf Grund der Definition einer stetigen Funktion fast selbstverständlich:

Lemma: Jede im Punkt $x_0 \in D$ differenzierbare Funktion $f: D \rightarrow \mathbb{R}$ ist dort auch stetig.

Beweis: Für kleine Werte von h ist

$$f(x_0 + h) = f(x_0) + hf'(x_0) + h\tilde{f}(h);$$

der Grenzwert $\lim_{h \rightarrow 0} f(x_0 + h)$ ist daher wegen der Konstanz von $f'(x_0)$ und der Stetigkeit von $\tilde{f}(h)$ gleich $f(x_0)$. ■

Die lokale Linearisierbarkeit einer differenzierbaren Funktion kann aber nicht nur für theoretische Sätze wie diesen ausgenutzt werden, sondern auch praktisch zur näherungsweise Berechnung von Funktionswerten: Wenn wir den Fehlerterm $h\tilde{f}(h)$ vernachlässigen, ist für kleine Werte von h

$$f(x+h) \approx f(x) + hf'(x).$$

Angenommen, wir wollen den Logarithmus von 1,1 näherungsweise berechnen. Wie wir wissen, ist $\log 1 = 0$, und die Ableitung von $\log x$ ist $1/x$. Daher ist

$$\log 1,1 \approx \log 1 + \frac{0,1}{1} = 0,1.$$

Verglichen mit dem tatsächlichen Wert bei ungefähr 0,09531 ist das, vor allem angesichts des geringen Rechenaufwands, keine schlechte Näherung. Auch wenn wir $1/99$ als über $1/100$ ausrechnen, erhalten wir mit

$$\frac{1}{99} = \frac{1}{100-1} \approx \frac{1}{100} + \frac{1}{100^2} = 0,0101$$

eine sehr gute Näherung für den exakten Wert $0,\overline{01}$.

Früher spielten solche Näherungsformeln eine große praktische Rolle, da man vor der Allgegenwart billiger Computer und Taschenrechner Funktionswerte meist aus Tabellen entnehmen mußte. Interpolationsformeln wie die obige gestatteten es dabei, eine größere Genauigkeit zu erzielen als die oft eher grobmaschige Tabelle hergab.

Die nächste Eigenschaft differenzierbarer Funktionen dürfte wohl allen Lesern aus der Schule bekannt sein: Die Bestimmung lokaler Extrema. Zunächst sei kurz an die Definition erinnert:

Definition: Wir sagen, eine Funktion $f: D \rightarrow \mathbb{R}$ auf einer Teilmenge $D \subseteq \mathbb{R}$ habe ein *lokales* $\left\{ \begin{array}{l} \text{Maximum} \\ \text{Minimum} \end{array} \right\}$ im Punkt $x_0 \in D$, wenn es ein $\varepsilon > 0$ gibt, so daß das Intervall $(x_0 - \varepsilon, x_0 + \varepsilon)$ in D liegt und

$f(x) \stackrel{\leq}{\geq} f(x_0)$ ist für alle x aus diesem Intervall. Lokale Maxima und Minima fassen wir unter dem Oberbegriff *lokale Extrema* zusammen.

Lemma: $f: D \rightarrow \mathbb{R}$ sei eine differenzierbare Funktion. Falls f im Punkt $x_0 \in D$ ein lokales Extremum hat, ist $f'(x_0) = 0$.

Beweis: Wir gehen aus von der Formel $f(x_0+h) = f(x_0) + hf'(x_0) + h\tilde{f}(h)$. Falls $f'(x_0) = a$ nicht verschwindet, ist in der Umgebung von x_0 die lineare Näherung $f(x_0) + hf'(x_0)$ auf einer der beiden Seiten echt größer als $f(x_0)$, auf der anderen echt kleiner. Wir müssen zeigen, daß dies auch für die Funktion f selbst gilt.

Dazu verwenden wir die Stetigkeit von \tilde{f} : Da $\tilde{f}(0) = 0$ ist, gibt es ein $\delta > 0$, so daß $|\tilde{f}(h)| < hf'(x_0)$ ist für alle h mit $|h| < \delta$. Ist $f'(x_0) > 0$, gelten daher für alle h mit $0 < h < \delta$ die beiden Formeln

$$f(x_0+h) = f(x_0) + hf'(x_0) + h\tilde{f}(h) > f(x_0) + hf'(x_0) - hf'(x_0) = f(x_0)$$

und

$$f(x_0-h) = f(x_0) - hf'(x_0) + h\tilde{f}(h) < f(x_0) - hf'(x_0) + hf'(x_0) = f(x_0).$$

Damit kann f in x_0 weder ein lokales Maximum noch ein lokales Minimum haben. Entsprechendes gilt für $f'(x_0) < 0$, wobei sich hier das Größer- und das Kleinerzeichen vertauschen. Somit ist ein lokales Extremum höchstens möglich, wenn $f'(x_0) = 0$ ist. ■

Natürlich *muß* die Funktion kein lokales Extremum in x_0 haben, wenn $f'(x_0) = 0$ ist: Das bekannteste Gegenbeispiel ist die Funktion $f(x) = x^3$, deren Ableitung $3x^2$ im Nullpunkt verschwindet, obwohl die Funktionswerte links der Null kleiner $f(0) = 0$ sind und rechts davon größer.

Satz von Rolle: $f: D \rightarrow \mathbb{R}$ sei eine differenzierbare Funktion, das abgeschlossene Intervall $[a, b]$ liege ganz in D , und $f(a) = f(b) = 0$. Dann gibt es ein $x_0 \in (a, b)$, so daß $f'(x_0)$ verschwindet; zwischen zwei Nullstellen der Funktion liegt also mindestens eine Nullstelle der Ableitung.

Beweis: Als differenzierbare Funktion ist f insbesondere stetig; wie wir aus dem vorigen Kapitel wissen, gibt es daher Punkte $x_m, x_M \in [a, b]$,

so daß $f(x_m)$ das Infimum und $f(x_M)$ das Supremum von f in $[a, b]$ ist. Sind beide Schranken gleich Null, ist die Funktion konstant und ihre Ableitung verschwindet überall. Falls mindestens eine der Schranken von Null verschieden ist, haben wir dort ein globales, insbesondere also lokales Extremum, und dort muß nach dem vorigen Lemma die Ableitung verschwinden. ■

Der französische Mathematiker MICHEL ROLLE (1652–1719) wurde in Ambert in der Auvergne als Sohn eines Kaufmanns geboren und erhielt nur eine rudimentäre Schulbildung. Er arbeitete als Gehilfe in verschiedenen Kanzleien in Ambert, bis er 1675 nach Paris zog. Auch dort arbeitete er vor allem als Schreiber und Rechner. Nebenbei brachte er sich im Selbststudium höhere Mathematik bei. Dies erlaubte ihm die Lösung eines damals populären mathematischen Problems, wofür ihm der Staatssekretär für die Marine eine Pension verschaffte. Außerdem wurde er dadurch in Mathematikerkreisen bekannt und bekam Kontakt zu anderen Mathematikern. Seine Forschungen beschäftigten sich vor allem mit Gleichungen höheren Grades, wo beispielsweise die Schreibweise $\sqrt[n]{x}$ auf ihn zurückgeht. Den Satz von ROLLE bewies er 1691, 1699 wurde er *Pensionnaire Géomètre* der Akademie der Wissenschaften.

Als wichtige Verallgemeinerung des Satzes von ROLLE erhalten wir den

Mittelwertsatz der Differentialrechnung: Ist $f: D \rightarrow \mathbb{R}$ eine differenzierbare Funktion und $[a, b] \subset D$, so gibt es ein $\xi \in (a, b)$, für das gilt:

$$f'(\xi) = \frac{f(b) - f(a)}{b - a}$$

Beweis: Wir betrachten die Funktion

$$g(x) = f(x) - \frac{f(b) - f(a)}{b - a}(x - a) - f(a).$$

Mit f ist auch g differenzierbar; außerdem ist $g(a) = f(a) - f(a) = 0$ und

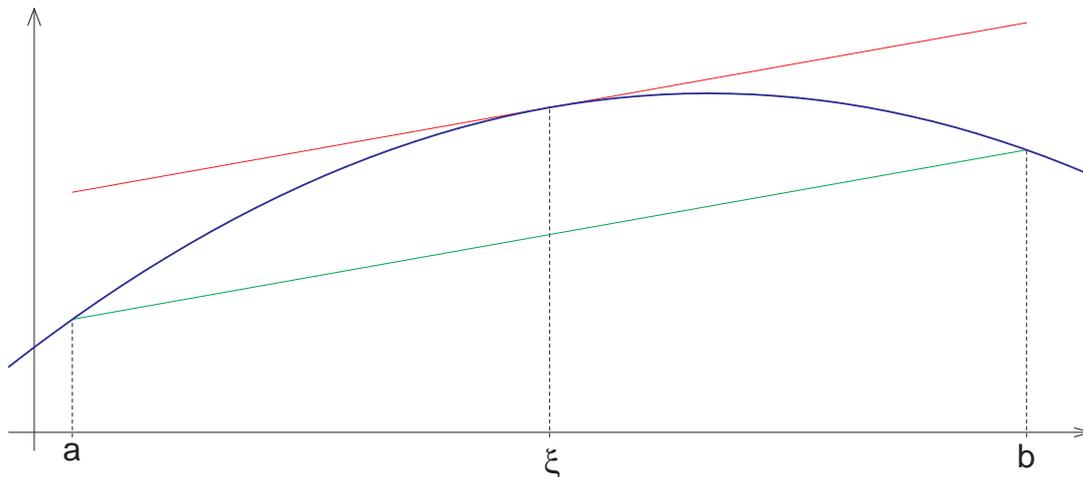
$$g(b) = f(b) - \frac{f(b) - f(a)}{b - a}(b - a) - f(a) = 0.$$

Damit können wir den Satz von ROLLE anwenden und bekommen ein $\xi \in (a, b)$ mit

$$g'(\xi) = f'(\xi) - \frac{f(b) - f(a)}{b - a} = 0.$$

Daraus folgt, daß $f'(\xi)$ den verlangten Wert hat. ■

Geometrisch bedeutet der Mittelwertsatz, daß es (mindestens) einen Punkt ξ zwischen a und b gibt, in dem die Tangente parallel ist zur Sehne durch die beiden Punkte $(a, f(a))$ und $(b, f(b))$:



An einem Zwischenpunkt ist die Tangente parallel zur Sehne

Wenn wir die Formel aus dem Mittelwertsatz nach $f(b)$ auflösen, erhalten wir $f(b) = f(a) + (b - a)f'(\xi)$. Setzen wir hierin $a = x_0$ und $b = x_0 + h$, wird diese Formel zu $f(x_0 + h) = f(x_0) + hf'(\xi)$. Dabei liegt ξ zwischen x_0 und $x_0 + h$, es gibt also ein $\eta \in (0, 1)$, so daß $\xi = x_0 + \eta h$ ist. Damit haben wir gezeigt

Mittelwertsatz, 2. Fassung: Ist $f: D \rightarrow \mathbb{R}$ eine differenzierbare Funktion und enthält D das Intervall von x_0 bis $x_0 + h$, so gibt es ein $\eta \in (0, 1)$, so daß gilt:

$$f(x_0 + h) = f(x_0) + hf'(x_0 + \eta h)$$

■

Der Mittelwertsatz ist eine eher technische Aussage, ist aber sehr nützlich, um viele interessante Aussagen zu beweisen. Beispielsweise wissen wir bereits, daß die Ableitung einer konstanten Funktion überall verschwindet; mit Hilfe des Mittelwertsatzes können wir auch die Umkehrung beweisen:

Lemma: Die Funktion $f: (a, b) \rightarrow \mathbb{R}$ sei differenzierbar und habe überall die Ableitung Null. Dann ist f konstant auf (a, b) .

Beweis: Wenn f nicht konstant wäre, gäbe es zwei Punkte $u, v \in (a, b)$ mit $f(u) \neq f(v)$. Nach dem Mittelwertsatz gäbe es dann ein ξ zwischen u und v , so daß

$$f'(\xi) = \frac{f(v) - f(u)}{v - u} \neq 0$$

wäre, im Widerspruch zur Voraussetzung. ■

Dieses Lemma kann falsch werden, wenn wir das offene Intervall (a, b) durch eine allgemeinere Menge ersetzen: Ist beispielsweise D die Vereinigung der offenen Intervalle $(0, 1)$ und $(2, 3)$ und $f(x) = 0$ auf $(0, 1)$, aber $f(x) = 1$ für $x \in (2, 3)$, so verschwindet $f'(x)$ auf ganz D , aber die Funktion ist nur auf den beiden Teilintervallen konstant, nicht auf ganz D . Der obige Beweis über den Mittelwertsatz ist in dieser Situation nicht anwendbar, da das Intervall $[u, v]$ nicht in D liegt.

Als nächste Anwendung können wir alle Funktionen mit konstanter Wachstumsrate bestimmen:

Lemma: Die Funktion $f: (a, b) \rightarrow \mathbb{R}$ sei differenzierbar und es gebe ein $c \in \mathbb{R}$, so daß $f'(x) = cf(x)$ für alle $x \in (a, b)$. Dann gibt es ein $y_0 \in \mathbb{R}$, so daß $f(x) = y_0 e^{cx}$ ist.

Beweis: Wir betrachten die Funktion $g(x) = e^{-cx} f(x)$. Als Produkt zweier differenzierbarer Funktionen ist sie differenzierbar auf ganz (a, b) und hat nach der Produktregel die Ableitung

$$g'(x) = -ce^{-cx} f(x) + e^{-cx} f'(x) = -ce^{-cx} f(x) + e^{-cx} \cdot cf(x) = 0.$$

Somit ist $g(x)$ konstant, es gibt also ein $y_0 \in \mathbb{R}$, so daß

$$g(x) = e^{-cx} f(x) = y_0$$

ist für alle $x \in (a, b)$. Durch Multiplikation mit e^{cx} folgt die Behauptung. ■

Als weitere Anwendung wollen wir die Regel von DE L'HÔPITAL betrachten: Die Gleichung

$$\frac{x^2 - 1}{x - 1} = x + 1$$

gilt streng genommen nicht für alle $x \in \mathbb{R}$, denn für $x = 1$ dividieren wir die Null durch die Null, was nicht erlaubt ist. Wir können allerdings sagen, daß

$$\lim_{x \rightarrow 1} \frac{x^2 - 1}{x - 1} = \lim_{x \rightarrow 1} x + 1 = 2$$

ist und in diesem Sinne die Funktion $\frac{x^2-1}{x-1}$ an ihrer „Definitionslücke“ stetig ergänzen.

Auch kompliziertere Brüche führen auf Ausdrücke der Form „0/0“, beispielsweise verschwinden für

$$\frac{e^x - 1}{x}$$

an der Stelle $x = 0$ sowohl der Zähler als auch der Nenner. Hier können wir nicht einfach kürzen, aber häufig können wir in solchen Fällen trotzdem Grenzwerte ausrechnen.

Wir gehen aus von zwei im Punkt x_0 differenzierbaren Funktionen f und g mit $f(x_0) = g(x_0)$, und wir interessieren uns für den Quotienten $f(x)/g(x)$. Für $x = x_0 + h$ in der Nähe von x_0 können wir schreiben

$$\frac{f(x)}{g(x)} = \frac{f(x_0 + h)}{g(x_0 + h)} = \frac{f(x_0) + hf'(x_0 + \eta_h h)}{g(x_0) + hg'(x_0 + \lambda_h h)} = \frac{f'(x_0 + \eta_h h)}{g'(x_0 + \lambda_h h)}$$

mit geeigneten Zahlen $\eta_h, \lambda_h \in (0, 1)$, die im allgemeinen von h abhängen. Geht h gegen Null, gehen aber auf jeden Fall auch $\eta_h h$ und $\lambda_h h$ gegen Null; damit folgt die

Regel von de l'Hôpital: Sind f, g im Punkt x_0 differenzierbar und ist $f(x_0) = g(x_0) = 0$, so ist

$$\lim_{x \rightarrow x_0} \frac{f(x)}{g(x)} = \lim_{x \rightarrow x_0} \frac{f'(x)}{g'(x)}.$$

■



GUILLAUME FRANÇOIS ANTOINE MARQUIS DE L'HÔPITAL, MARQUIS DE SAINTE-MESME, COMTE D'ENTREMONT ET SEIGNEUR D'OUQUES-LA-CHAISE (1661–1704) interessierte sich seit seiner Kindheit vor allem für Mathematik, schlug aber, der Familientradition entsprechend, eine militärische Karriere ein. Wegen seiner Kurzsichtigkeit gab er diese später auf und widmete sich ganz der Mathematik, wobei er zeitweise JOHANN BERNOULLI (1667–1748) als Privatlehrer engagierte. Gestützt auf dessen Vorlesung veröffentlichte er 1696 das erste Analysis-Lehrbuch. Es war sehr erfolgreich und erschien bis 1781 in immer neuen Auflagen.

Tatsächlich muß man bei der Regel von DE L'HÔPITAL nicht voraussetzen, daß die Funktionen f und g sowie ihre Ableitungen im Punkt x_0 definiert sind: Wir können auch ausgehen von Funktionen, die auf einem offenen Intervall (a, b) definiert sind und dann die Grenzwerte bei Annäherung an die Intervallgrenzen betrachten. Da die Funktionen für $x \leq a$ und $x \geq b$ nicht definiert sind, können wir dazu freilich nur *einseitige* Grenzwerte betrachten. Diese sind folgendermaßen definiert:

Definition: Für eine Funktion $f: (a, b) \rightarrow \mathbb{R}$ ist

$$\lim_{x \nearrow b} f(x) = c,$$

wenn es zu jedem $\varepsilon > 0$ ein $\delta > 0$ gibt, so daß $|c - f(x)| < \varepsilon$ ist für alle $x \in (b - \delta, b)$. Wir sagen, dieser Grenzwert sei ∞ , wenn es zu jedem $M \in \mathbb{R}$ ein $\delta > 0$ gibt, so daß $f(x) > M$ für alle $x \in (b - \delta, b)$, und er ist $-\infty$, falls es zu jedem $M \in \mathbb{R}$ ein $\delta > 0$ gibt, so daß $f(x) < M$ für alle $x \in (b - \delta, b)$. Entsprechend ist

$$\lim_{x \searrow a} f(x) = c,$$

wenn es zu jedem $\varepsilon > 0$ ein $\delta > 0$ gibt, so daß $|c - f(x)| < \varepsilon$ ist für alle $x \in (a, a + \delta)$. Wir sagen, dieser Grenzwert sei ∞ , wenn es zu jedem $M \in \mathbb{R}$ ein $\delta > 0$ gibt, so daß $f(x) > M$ für alle $x \in (a, a + \delta)$, und er ist $-\infty$, falls es zu jedem $M \in \mathbb{R}$ ein $\delta > 0$ gibt, so daß $f(x) < M$ für alle $x \in (a, a + \delta)$.

Diese Definitionen sollen auch gelten für $a = -\infty$ und/oder $b = \infty$, wobei in diesem Fall $(-\infty, \infty) = \mathbb{R}$ sein soll, (a, ∞) für $a \neq -\infty$ die Menge aller reeller Zahlen größer a und $(-\infty, b)$ für $b \neq \infty$ die

Menge aller reeller Zahlen kleiner b . Anstelle des Intervalls $(b - \delta, b)$ muß dann ein Intervall (N, ∞) betrachtet werden, anstelle von $(a, a + \delta)$ entsprechend $(-\infty, N)$.

Regel von de l'Hôpital, 2. Fassung: Die Funktionen $f, g: (a, b)$ seien differenzierbar, $g(x) \neq 0$ und $g'(x) \neq 0$ für alle $x \in (a, b)$.

a) Falls $\lim_{x \nearrow b} f(x) = \lim_{x \nearrow b} g(x) = 0$ ist und $\lim_{x \nearrow b} \frac{f'(x)}{g'(x)}$ existiert, ist

$$\lim_{x \nearrow b} \frac{f(x)}{g(x)} = \lim_{x \nearrow b} \frac{f'(x)}{g'(x)}.$$

b) Falls $\lim_{x \searrow a} f(x) = \lim_{x \searrow a} g(x) = 0$ ist und $\lim_{x \searrow a} \frac{f'(x)}{g'(x)}$ existiert, ist

$$\lim_{x \searrow a} \frac{f(x)}{g(x)} = \lim_{x \searrow a} \frac{f'(x)}{g'(x)}.$$

c) Ist $\lim_{x \nearrow b} g(x) = \pm\infty$ und existiert $\lim_{x \nearrow b} \frac{f'(x)}{g'(x)}$, so ist

$$\lim_{x \nearrow b} \frac{f(x)}{g(x)} = \lim_{x \nearrow b} \frac{f'(x)}{g'(x)}.$$

d) Ist $\lim_{x \searrow a} g(x) = \pm\infty$ und existiert $\lim_{x \searrow a} \frac{f'(x)}{g'(x)}$, so ist

$$\lim_{x \searrow a} \frac{f(x)}{g(x)} = \lim_{x \searrow a} \frac{f'(x)}{g'(x)}.$$

Beweis: Wir wollen den Quotienten $f'(x)/g'(x)$ umschreiben als Ableitung einer Funktion $q(x)$. Die Gleichung

$$q'(x) = \frac{f'(x)}{g'(x)} = f'(x) \cdot \frac{1}{g'(x)}$$

kann beispielsweise so zustande kommen, daß wir die Kettenregel anwenden auf die Hintereinanderausführung von f und einer Funktion mit Ableitung $1/g'(x)$; eine solche Funktion ist die Umkehrfunktion g^{-1} von g . Um q definieren zu können, müssen wir uns somit als erstes überlegen, ob und wo diese Umkehrfunktion existiert.

Da wir vorausgesetzt haben, daß $g'(x)$ nirgends verschwindet, ist g eine injektive Abbildung: Hätte nämlich ein Punkt $c \in \mathbb{R}$ zwei verschiedene Urbilder $x_1, x_2 \in (a, b)$, so könnten wir den Satz von ROLLE auf die Funktion $g(x) - c$ anwenden und erhielten zwischen x_1 und x_2 eine Nullstelle x_0 der Ableitung $g'(x)$. Da g als differenzierbare Funktion insbesondere stetig ist, können wir daher den Satz aus dem vorigen Kapitel anwenden, wonach eine auf einem Intervall injektive Funktion dort streng monoton ist. Indem wir nötigenfalls f durch $-f$ und g durch $-g$ ersetzen, können wir annehmen, daß g streng monoton wächst. Setzen wir

$$c = \lim_{x \searrow a} g(x) \quad \text{und} \quad d = \lim_{x \nearrow b} g(x),$$

bildet g dann das Intervall (a, b) ab auf (c, d) und hat eine differenzierbare Umkehrfunktion $g^{-1}: (c, d) \rightarrow (a, b)$. Diese können wir schachteln mit der Zählerfunktion f zu einer Funktion $q = f \circ g^{-1}: (c, d) \rightarrow \mathbb{R}$, die einen Punkt $y \in (c, d)$ abbildet auf $f(x)$, wobei $x \in (a, b)$ das Urbild von y unter g ist. Nach der Kettenregel und der Regel über die Ableitung der Umkehrfunktion ist, wie gewünscht,

$$q'(y) = f'(x) \cdot \frac{1}{g'(x)} = \frac{f'(x)}{g'(x)}.$$

Sofern diese Grenzwerte existieren, ist daher

$$\lim_{x \nearrow b} \frac{f'(x)}{g'(x)} = \lim_{y \nearrow d} q'(y) \quad \text{und} \quad \lim_{x \searrow a} \frac{f'(x)}{g'(x)} = \lim_{y \searrow c} q'(y).$$

Auch den Grenzwert von $f(x)/g(x)$ können wir mit Hilfe von q ausdrücken: Mit $y = g(x)$ ist,

$$\frac{f(x)}{g(x)} = \frac{f(g^{-1}(y))}{y} = \frac{q(y)}{y},$$

also ist, sofern die Grenzwerte existieren,

$$\lim_{x \nearrow b} \frac{f(x)}{g(x)} = \lim_{y \nearrow d} \frac{q(y)}{y} \quad \text{und} \quad \lim_{x \searrow a} \frac{f(x)}{g(x)} = \lim_{y \searrow c} \frac{q(y)}{y}.$$

Wir müssen somit für $a)$ und $c)$ zeigen, daß

$$\lim_{y \nearrow d} q'(y) = \lim_{y \nearrow d} \frac{q(y)}{y}$$

ist und für $b)$ und $d)$ entsprechend die Gleichheit

$$\lim_{y \searrow c} q'(y) = \lim_{y \searrow c} \frac{q(y)}{y}.$$

$a)$ Hier ist, da wir g als monoton wachsend vorausgesetzt haben, $d = 0$, also müssen wir zeigen, daß für eine in einem Intervall $(c, 0)$ differenzierbare Funktion q mit $\lim_{y \nearrow 0} q(y) = 0$ gilt

$$\lim_{y \nearrow 0} q'(y) = \lim_{y \nearrow 0} \frac{q(y)}{y}.$$

Wir betrachten zunächst den Fall, daß der links stehende Grenzwert verschwindet.

Nach Definition des Grenzwerts gibt es dann zu jedem $\varepsilon > 0$ ein $\delta > 0$, so daß $|q'(y)| < \varepsilon$ ist für alle $y \in (-\delta, 0)$. Ist y ein beliebiger Punkt aus diesem Intervall und $z > y$, so gibt es nach dem Mittelwertsatz ein $\xi \in (y, z) \subset (-\delta, 0)$ so daß

$$\frac{q(z) - q(y)}{z - y} = q'(\xi)$$

ist und damit

$$|q(y) - q(z)| = |q'(\xi)| |y - z| < \varepsilon \cdot |y - z|.$$

Setzen wir hier für z die Glieder einer Nullfolge aus $(-\delta, 0)$ ein, erhalten wir für den Grenzwert die Ungleichung

$$|q(y)| = \left| q(y) - \lim_{z \nearrow 0} q(z) \right| \leq \varepsilon |y|.$$

Somit ist

$$\left| \frac{q(y)}{y} \right| \leq \varepsilon \quad \text{für } -\delta < x < 0.$$

Dies zeigt, daß der Grenzwert für $y \rightarrow 0$ verschwinden muß.

Damit ist $a)$ bewiesen für den Fall, daß $\lim_{y \nearrow 0} q'(y)$ verschwindet. Ist dieser Limes eine von Null verschiedene Zahl m , so betrachten wir die neue Funktion $p(y) = q(y) - my$. Da $p'(y) = q'(y) - m$ ist, verschwindet hier der Grenzwert von $p'(y)$, also ist nach der gerade bewiesenen Formel

$$\lim_{y \nearrow 0} \frac{p(y)}{y} = \lim_{y \nearrow 0} \frac{q(y) - my}{y} = \lim_{y \nearrow 0} \frac{q(y)}{y} - m = 0,$$

womit auch hier die Behauptung folgt.

b) Hier ist entsprechend $c = 0$, das Definitionsintervall von q ist also $(0, d)$. Die Situation ist spiegelbildlich zu a); wenn wir die dort bewiesene Formel auf die Funktion $q(-x)$ anwenden, folgt die hier benötigte Formel

$$\lim_{y \searrow 0} q'(y) = \lim_{y \searrow 0} \frac{q(y)}{y}.$$

c) Hier ist $d = \infty$, wir müssen also zeigen, daß gilt

$$\lim_{y \rightarrow \infty} q'(y) = \lim_{y \rightarrow \infty} \frac{q(y)}{y}.$$

Wie in a) betrachten wir zunächst den Fall, daß der linksstehende Grenzwert verschwindet; der allgemeine Fall folgt daraus wie dort.

Nach Definition des Grenzwerts gibt es zu jedem $\varepsilon > 0$ ein $M \in \mathbb{R}$, so daß $|q'(y)| < \varepsilon/2$ ist für alle $y \geq M$, wobei wir für M eine positive reelle Zahl wählen können. Zu jedem $y_0 > M$ und jedem $y > y_0$ gibt es nach dem Mittelwertsatz ein $\eta > M$, so daß

$$\frac{q(y) - q(y_0)}{y - y_0} = q'(\eta)$$

ist und damit

$$|q(y) - q(y_0)| = |q'(\eta)| \cdot (y - y_0) < \frac{\varepsilon}{2}(y - y_0) < \frac{\varepsilon}{2}y$$

Nach der Dreiecksungleichung gilt somit

$$\left| \frac{q(y)}{y} \right| \leq \left| \frac{q(y) - q(y_0)}{y} \right| + \left| \frac{q(y_0)}{y} \right| < \frac{\varepsilon}{2} + \left| \frac{q(y_0)}{y} \right|.$$

$q(y_0)$ ist ein fester Wert; wir können daher ein $N \geq y_0$ finden, so daß $|q(y_0)/y| < \frac{1}{2}\varepsilon$ ist für $y > N$. Zu jedem $\varepsilon > 0$ gibt es daher ein $N \in \mathbb{R}$, so daß für alle $y > N$ gilt: $|q(y)/y| < \varepsilon$. Somit muß der Grenzwert für $y \rightarrow \infty$ die Null sein.

d) Hier ist entsprechend $c = -\infty$, das Definitionsintervall von q ist also $(-\infty, d)$. Die Situation ist spiegelbildlich zu c); wenn wir die dort bewiesene Formel auf die Funktion $q(-y)$ anwenden, folgt die hier benötigte Formel

$$\lim_{y \rightarrow -\infty} q'(y) = \lim_{y \rightarrow \infty} \frac{q(y)}{y}.$$

Damit sind alle Fälle vollständig bewiesen. ■

Die Regel von DE L'HÔPITAL kann auch dazu verwendet werden, um Grenzwerte der Form „ $0 \cdot \infty$ “ oder „ 0^0 “ zu berechnen: Ist etwa $f: D \rightarrow \mathbb{R}$ eine differenzierbare Funktion, deren Grenzwert für $x \rightarrow x_0$ verschwindet und ist $g: D \setminus \{x_0\} \rightarrow \mathbb{R}$ eine differenzierbare Funktion, die für $x \rightarrow x_0$ unbeschränkt wächst, so können wir auf der Teilmenge von $D \setminus \{x_0\}$, auf der g keine Nullstellen hat, die Funktion $1/g$ betrachten. Da g für $x \rightarrow x_0$ unbeschränkt wächst, geht $1/g$ dann gegen Null; schreiben wir also

$$f(x) \cdot g(x) = \frac{f(x)}{1/g(x)},$$

sind wir in der gerade betrachteten Situation und können die Regel von DE L'HÔPITAL anwenden. Genauso können wir auch

$$f(x) \cdot g(x) = \frac{g(x)}{1/f(x)}$$

schreiben und sind dann in der Situation „ ∞/∞ “. Welche der beiden Vorgehensweisen nützlicher ist, hängt von den jeweiligen Funktionen f und g ab.

Bei Grenzwerten der Form „ 0^0 “ haben wir Funktionen der Form $f(x)^{g(x)}$. Da wir allgemeine Potenzen a^x nur für $a > 0$ definiert haben, müssen wir hier ausgehen von einer differenzierbaren Funktion $f: D \rightarrow \mathbb{R}$, die für alle $x \neq x_0$ aus D positive Werte annimmt, deren Grenzwert für $x \rightarrow x_0$ aber verschwindet. Von der differenzierbaren Funktion $g: D \rightarrow \mathbb{R}$ setzen wir nur voraus, daß $g(x_0)$ verschwindet. Dann ist für $x \neq x_0$

$$f(x)^{g(x)} = e^{g(x) \cdot \log f(x)}.$$

Im Exponenten steht für $x \rightarrow x_0$ ein Ausdruck der Form „ $0 \cdot \infty$ “; wenn wir dafür den Grenzwert für $x \rightarrow x_0$ berechnen können, kennen wir wegen der Stetigkeit der Exponentialfunktion auch den für $f(x)^{g(x)}$.

Als Beispiel betrachten wir $\lim_{x \searrow 0} x^x$. Da $x^x = e^{x \log x}$ ist, müssen wir als erstes den Grenzwert für $x \log x$ berechnen:

$$\lim_{x \searrow 0} x \log x = \lim_{x \searrow 0} \frac{\log x}{1/x} = \lim_{x \searrow 0} \frac{1/x}{-1/x^2} = \lim_{x \searrow 0} (-x) = 0.$$

Der gesuchte Grenzwert ist somit $e^0 = 1$.

Als vorläufig letzte Anwendung des Mittelwertsatzes wollen noch einen anschaulich fast klaren Zusammenhang zwischen Ableitung und Monotonie festhalten:

Lemma: $f: (a, b) \rightarrow \mathbb{R}$ sei eine differenzierbare Funktion.

- a) Ist $f'(x) \geq 0$ für alle $x \in (a, b)$, so ist f monoton wachsend.
- b) Ist $f'(x) > 0$ für alle $x \in (a, b)$, so ist f streng monoton wachsend.
- c) Ist $f'(x) \leq 0$ für alle $x \in (a, b)$, so ist f monoton fallend.
- d) Ist $f'(x) < 0$ für alle $x \in (a, b)$, so ist f streng monoton fallend.

Beweis: a) Gäbe es in (a, b) zwei Punkte $x_1 < x_2$ mit $f(x_1) > f(x_2)$, so gäbe es nach dem Mittelwertsatz ein $\xi \in (x_1, x_2)$ mit

$$f'(\xi) = \frac{f(x_2) - f(x_1)}{x_2 - x_1} < 0,$$

im Widerspruch zur Voraussetzung.

b) Gäbe es in (a, b) zwei Punkte $x_1 < x_2$ mit $f(x_1) \geq f(x_2)$, so gäbe es nach dem Mittelwertsatz ein $\xi \in (x_1, x_2)$ mit

$$f'(\xi) = \frac{f(x_2) - f(x_1)}{x_2 - x_1} \leq 0,$$

im Widerspruch zur Voraussetzung.

c) und d) folgen entweder genauso, oder indem man a) und b) anwendet auf die Funktion $-f$. ■

Die Umkehrungen von b) und d) gelten nicht: Beispielsweise ist die Funktion $f(x) = x^3$ streng monoton wachsend, aber trotzdem verschwindet ihre Ableitung im Punkt Null.

Zum Abschluß dieses Paragraphen wollen wir den Mittelwertsatz noch etwas verallgemeinern zu einer Aussage, die im nächsten Paragraphen sehr nützlich sein wird:

Verallgemeinerter Mittelwertsatz der Differentialrechnung: Die beiden Funktionen $f, g: (a, b) \rightarrow \mathbb{R}$ seien differenzierbar auf (a, b) , und g' habe dort keine Nullstelle. Dann gibt es zu jedem $x \in (a, b)$ und jedem

$h \in \mathbb{R} \setminus \{0\}$, für das auch noch $x + h$ in (a, b) liegt, eine reelle Zahl $0 < \eta < 1$, so daß gilt:

$$\frac{f(x+h) - f(x)}{g(x+h) - g(x)} = \frac{f'(x+\eta h)}{g'(x+\eta h)}.$$

Zum *Beweis* betrachten wir die Funktion $\varphi: [0, 1] \rightarrow \mathbb{R}$ mit

$$\varphi(\eta) = f(x+\eta h)(g(x+h) - g(x)) - g(x+\eta h)(f(x+h) - f(x)).$$

Einsetzen zeigt, daß $\varphi(0) = \varphi(1) = f(a)g(b) - f(b)g(a)$ ist, die Funktion $\varphi(\eta) - \varphi(0)$ verschwindet also an beiden Intervallenden. Nach dem Satz von ROLLE gibt es daher einen Punkt $\eta \in (0, 1)$, in dem ihre Ableitung $\varphi'(\eta)$ verschwindet. Nach der Kettenregel ist

$$\varphi'(\eta) = h \left(f'(x+\eta h)(g(x+h) - g(x)) - g'(x+\eta h)(f(x+h) - f(x)) \right)$$

und h ist nach Voraussetzung von Null verschieden, also verschwindet die Klammer. Die Differenz $g(x+h) - g(x)$ kann nicht verschwinden, denn sonst müßte es nach dem gleichen Argument wie eben zwischen x und $x+h$ eine Nullstelle von g' geben, was im Satz ausgeschlossen wurde. Aus diesem Grund ist auch $g'(x+\eta h) \neq 0$. Somit können wir dividieren und erhalten die Gleichung

$$\frac{f(x+h) - f(x)}{g(x+h) - g(x)} = \frac{f'(x+\eta h)}{g'(x+\eta h)}.$$

■

§3: Höhere Ableitungen und Taylor-Entwicklung

Wenn die Funktion $f: D \rightarrow \mathbb{R}$ auf $D \subseteq \mathbb{R}$ differenzierbar ist, definieren die Werte der Ableitungen eine neue Funktion $f': D \rightarrow \mathbb{R}$. Diese kann wieder differenzierbar sein, muß es aber nicht: Die Funktion $f: \mathbb{R} \rightarrow \mathbb{R}$ mit

$$f(x) = \begin{cases} x^2 & \text{falls } x \geq 0 \\ -x^2 & \text{sonst} \end{cases}$$

etwa ist für alle $x_0 \in \mathbb{R}$ differenzierbar: Der einzige problematische Wert ist $x_0 = 0$, und dort haben x^2 und $-x^2$ dieselbe Ableitung, die somit gleich der von f ist. Für $x_0 > 0$ ist $f'(x_0) = 2x_0$, für $x < 0$ ist $f'(x_0) = -2x_0$ und für $x_0 = 0$ ist $f'(x_0) = 0$. Zusammenfassend können wir also sagen, daß $f'(x) = |x|$ ist, und von dieser Funktion wissen wir, daß sie im Nullpunkt nicht differenzierbar ist.

Betrachten wir dagegen die Funktion $f(x) = x^2$, so ist $f'(x) = 2x$ als lineare Funktion wieder differenzierbar und hat als Ableitung die konstante Funktion mit Wert zwei.

Falls f' eine differenzierbare Funktion ist, bezeichnen wir deren Ableitung als die *zweite Ableitung* f'' von f ; falls diese ebenfalls differenzierbar ist, heißt ihre Ableitung die *dritte Ableitung* f''' von f , und so weiter.

Formal ausgedrückt heißt das:

Definition: $f: D \rightarrow \mathbb{R}$ sei eine Funktion auf $D \subseteq \mathbb{R}$.

- a) f heißt (mindestens) k -fach differenzierbar, wenn f differenzierbar ist, und wenn für $k > 1$ die Funktion f' mindestens $(k - 1)$ -fach differenzierbar ist.
- b) Ist $f: D \rightarrow \mathbb{R}$ eine mindestens k -fach differenzierbare Funktion, so definieren wir ihre k -te Ableitung $f^{(k)}(x)$ im Falle $k = 1$ als $f^{(1)}(x) = f'(x)$ und für $k > 1$ als $f^{(k)}(x) = (f^{(k-1)})'(x)$.
- c) Eine k -fach differenzierbare Funktion heißt k -fach stetig differenzierbar, wenn $f^{(k)}$ eine stetige Funktion ist.

Eine k -fach differenzierbare Funktion muß nicht auch k -fach stetig differenzierbar sein; schon für eine differenzierbare Funktion muß die Ableitung keine stetige Funktion sein. Mit den Funktionen, die wir bislang kennen, lassen sich leider noch keine einfachen Beispiele dafür konstruieren, aber im nächsten Paragraphen werden wir welche kennenlernen.

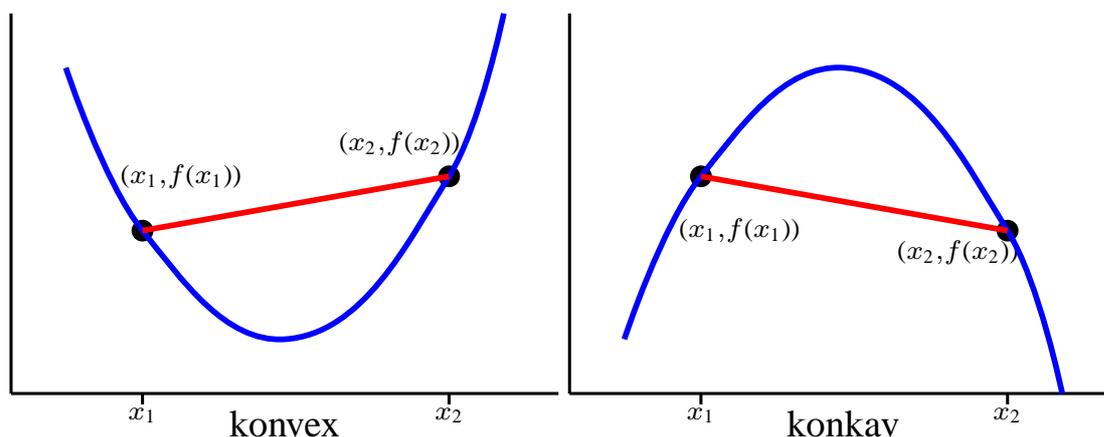
Was die Bezeichnung der höheren Ableitungen betrifft, so verwenden wir für kleine Werte von ℓ auch Striche:

$$f''(x) = f^{(2)}(x), \quad f'''(x) = f^{(3)}(x) \quad \text{und} \quad f''''(x) = f^{(4)}(x).$$

Die meisten unserer gängigen Funktionen sind beliebig oft differenzierbar: Die Ableitung eines Polynoms ist wieder ein Polynom und damit wieder differenzierbar, die Ableitung der Exponentialfunktion ist die Exponentialfunktion selbst, und auch bei rationalen Funktionen ist die Ableitung wieder eine rationale Funktion auf demselben Definitionsbereich und somit differenzierbar. Trotzdem gibt es natürlich Funktionen, die *genau* k -fach differenzierbar sind, beispielsweise die Funktion $f(x) = x^k |x|$ im Punkt $x = 0$.

Aus der zweiten Ableitung können wir, wie aus der ersten, auch geometrische Information über den Graphen der Funktion gewinnen. Während uns die erste Ableitung sagt, ob die Funktion in der Umgebung eines Punktes steigt oder fällt, sagt uns die zweite Ableitung etwas darüber, ob die Funktion konkav oder konvex ist.

Definition: Eine Funktion $f: D \rightarrow \mathbb{R}$ heißt *konvex* über dem Intervall $(a, b) \subseteq D$, wenn für alle $x_1, x_2 \in (a, b)$ der Graph von f über (a, b) unterhalb der Verbindungsstrecke der beiden Punkte $(x_1, f(x_1))$ und $(x_2, f(x_2))$ liegt; liegt er stets oberhalb, bezeichnen wir die Funktion als *konkav*.



Die lineare Abbildung $\varphi: \mathbb{R} \rightarrow \mathbb{R}$ mit

$$\varphi(\lambda) = (1 - \lambda)x_1 + \lambda x_2 = \lambda(x_2 - x_1) + x_1$$

bildet das offene Intervall $(0, 1)$ bijektiv ab auf das Intervall (x_1, x_2) , denn da $x_2 - x_1$ positiv ist, ist sie streng monoton wachsend, und sie bildet die Null ab auf x_1 und die Eins auf x_2 . Die Punkte zwischen x_1 und x_2 sind daher genau die Punkte der Form $(1 - \lambda)x_1 + \lambda x_2$ mit $\lambda \in (0, 1)$.

Die Gerade durch $(x_1, f(x_1))$ und $(x_2, f(x_2))$ hat die Gleichung

$$y = f(x_1) + \frac{f(x_2) - f(x_1)}{x_2 - x_1}(x - x_1);$$

für $x = (1 - \lambda)x_1 + \lambda x_2$ haben wir also die y -Koordinate

$$\begin{aligned} f(x_1) + \frac{f(x_2) - f(x_1)}{x_2 - x_1} ((1 - \lambda)x_1 + \lambda x_2 - x_1) \\ = f(x_1) + \frac{f(x_2) - f(x_1)}{x_2 - x_1} \cdot \lambda(x_2 - x_1) \\ = f(x_1) + \lambda(f(x_2) - f(x_1)) = (1 - \lambda)f(x_1) + \lambda f(x_2). \end{aligned}$$

In Formeln ausgedrückt ist die Funktion $f: D \rightarrow \mathbb{R}$ daher konvex über (a, b) , wenn

$$f((1 - \lambda)x_1 + \lambda x_2) \leq (1 - \lambda)f(x_1) + \lambda f(x_2)$$

ist für alle $x_1, x_2 \in (a, b)$ und alle $\lambda \in (0, 1)$; sie ist konkav, wenn die entsprechende Ungleichung mit \geq gilt.

Lemma: Eine zweimal stetig differenzierbare Funktion $f: D \rightarrow \mathbb{R}$ ist genau dann konvex über dem Intervall $(a, b) \subseteq D$, wenn $f''(x) \geq 0$ ist für alle $x \in (a, b)$; sie ist genau dann konkav über (a, b) , wenn $f''(x) \leq 0$ ist für alle $x \in (a, b)$.

Beweis: Wir zeigen zunächst, daß im Falle der Konvexität die zweite Ableitung in ganz (a, b) größer oder gleich null sein muß: Andernfalls gäbe es ein $x_0 \in (a, b)$ mit $f''(x_0) < 0$. Wir betrachten die Funktion $g(x) \stackrel{\text{def}}{=} f(x) - f'(x_0)(x - x_0)$. Als Summe von f und einer linearen Funktion ist g zweimal differenzierbar mit

$$g'(x_0) = f'(x_0) - f'(x_0) = 0 \quad \text{und} \quad g''(x_0) = f''(x_0) < 0.$$

Die Funktion $g'(x)$ ist also in einem hinreichend kleinen Intervall $(x_0 - h, x_0 + h)$ streng monoton fallend; sie ist daher positiv für $x < x_0$ und negativ für $x > x_0$. Somit ist g streng monoton wachsend für $x < x_0$ und streng monoton fallend für $x > x_0$; die Funktion g hat also bei x_0 ein lokales Maximum. Für ein $\varepsilon < h$ ist daher $g(x_0 \pm \varepsilon) < g(x_0)$ und damit ist auch

$$f(x_0) = g(x_0) > \frac{1}{2}g(x_0 - \varepsilon) + \frac{1}{2}g(x_0 + \varepsilon) = \frac{1}{2}f(x_0 - \varepsilon) + \frac{1}{2}f(x_0 + \varepsilon).$$

Dies widerspricht aber der Konvexitätsbedingung für $x_{1/2} = x_0 \pm \varepsilon$ und $\lambda = \frac{1}{2}$. Somit muß $f''(x)$ in ganz (a, b) größer oder gleich null sein.

Umgekehrt sei $f''(x) \geq 0$ für alle $x \in (a, b)$; wir müssen zeigen, daß f dann konvex ist. Seien also $x_1 < x_2$ zwei beliebige Punkte aus (a, b) und $x = (1 - \lambda)x_1 + \lambda x_2$ mit $\lambda \in (0, 1)$. Nach dem Mittelwertsatz gibt es Punkte $\xi_1 \in (x_1, x)$ und $\xi_2 \in (x, x_2)$, so daß

$$f'(\xi_1) = \frac{f(x) - f(x_1)}{x - x_1} = \frac{f(x) - f(x_1)}{\lambda(x_2 - x_1)} \quad \text{und}$$

$$f'(\xi_2) = \frac{f(x_2) - f(x)}{x_2 - x} = \frac{f(x_2) - f(x)}{(1 - \lambda)(x_2 - x_1)}$$

ist. Da f'' nirgends negativ wird, ist f' monoton wachsend und damit insbesondere $f'(\xi_1) \leq f'(\xi_2)$. Diese Ungleichung können wir auch schreiben als

$$\frac{f(x) - f(x_1)}{\lambda(x_2 - x_1)} \leq \frac{f(x_2) - f(x)}{(1 - \lambda)(x_2 - x_1)},$$

und da $x_1 < x_2$ ist, folgt daraus

$$\frac{f(x) - f(x_1)}{\lambda} \leq \frac{f(x_2) - f(x)}{1 - \lambda}.$$

Für $\lambda \in (0, 1)$ ändert sich nichts an dieser Ungleichung, wenn wir mit $\lambda(1 - \lambda)$ multiplizieren; dies führt auf

$$(1 - \lambda)f(x) - (1 - \lambda)f(x_1) \leq \lambda f(x_2) - \lambda f(x)$$

und damit die gewünschte Ungleichung

$$f(x) \leq (1 - \lambda)f(x_1) + \lambda f(x_2),$$

die die Konvexität von f ausdrückt. Damit ist die Behauptung für konvexe Funktionen bewiesen.

Für konkave Funktionen folgt sie einfach daraus, daß $-f$ für eine konkave Funktion f konvex ist. ■

Eine der wichtigsten Anwendungen mehrfach differenzierbarer Funktionen ist eine Formel, die in vielen Fällen zur Berechnung beliebig genauer Näherungswerte führt, die TAYLOR-Formel:

Satz: $f: (a, b) \rightarrow \mathbb{R}$ sei mindestens $(n + 1)$ -fach stetig differenzierbar auf dem Intervall $(a, b) \subseteq \mathbb{R}$. Dann gilt für jedes x aus (a, b) und jedes

$h \in \mathbb{R}$ mit $x + h \in (a, b)$ die Formel

$$\begin{aligned} f(x+h) &= f(x) + hf'(x) + \frac{h^2}{2}f''(x) + \cdots + \frac{h^n}{n!}f^{(n)}(x) + R_{n+1}(x, h) \\ &= \sum_{k=0}^n \frac{h^k}{k!}f^{(k)}(x) + R_{n+1}(x, h) \end{aligned}$$

mit dem Restglied $R_{n+1}(x, h) = \frac{h^{n+1}}{(n+1)!}f^{(n+1)}(x + \eta h)$ zu einer reellen Zahl η zwischen 0 und 1.

Beweis: Wir betrachten für festgehaltenes x und h die Hilfsfunktion

$$F: \begin{cases} (a, b) \rightarrow \mathbb{R} \\ y \mapsto \sum_{k=0}^n \frac{f^{(k)}(y)}{k!}(x+h-y)^k \end{cases} .$$

Für $y = x$ ist $x + h - y = h$ und

$$F(x) = \sum_{k=0}^n \frac{f^{(k)}(x)}{k!}h^k ;$$

dieses Polynom wollen wir kurz mit $T_{f,x,n}(h)$ bezeichnen. Für $y = x+h$ ist $x + h - y = 0$, so daß alle Terme mit positivem k verschwinden und nur der konstante Term $F(x+h) = f(x+h)$ übrig bleibt. Wenn die behauptete Formel stimmt, ist also $F(x+h) - F(x)$ gleich dem Restglied.

Die Ableitung von F ist nach der Produkt- und Kettenregel

$$\begin{aligned} F'(y) &= \sum_{k=0}^n \frac{f^{(k+1)}(y)}{k!}(x+h-y)^k - \sum_{k=1}^n \frac{f^{(k)}(y)}{k!} \cdot k \cdot (x+h-y)^{k-1} \\ &= \sum_{k=1}^{n+1} \frac{f^{(k)}(y)}{(k-1)!}(x+h-y)^{k-1} - \sum_{k=1}^n \frac{f^{(k)}(y)}{(k-1)!} \cdot (x+h-y)^{k-1} \\ &= \frac{f^{(n+1)}(y)}{n!}(x+h-y)^n . \end{aligned}$$

Nach dem verallgemeinerten Mittelwertsatz der Differentialrechnung gibt es für jede zwischen x und $x+h$ differenzierbare Funktion g eine

reelle Zahl η zwischen 0 und 1, so daß gilt

$$\frac{F'(x + \eta h)}{g'(x + \eta h)} = \frac{F(x + h) - F(x)}{g(x + h) - g(x)} = \frac{f(x + h) - T_{f,x,n}(x)}{g(x + h) - g(x)}.$$

Diese Gleichung können wir nach $f(x + h)$ auflösen als

$$\begin{aligned} f(x + h) &= T_{f,x,n}(h) + \frac{g(x + h) - g(x)}{g'(x + \eta h)} F'(x + \eta h) \\ &= T_{f,x,n}(h) + \frac{g(x + h) - g(x)}{g'(x + \eta h)} \frac{f^{(n+1)}(x + \eta h)}{n!} (h - \eta h)^n, \end{aligned}$$

wobei wir in der zweiten Zeile die obige Formel für $F'(y)$ benutzt haben.

Speziell für die Funktion $g(y) = (x + h - y)^{n+1}$ ist $g(x + h) = 0$, $g(x) = h^{n+1}$ und $g'(x + \eta h) = -(n + 1)(h - \eta h)^n$, wir erhalten also

$$\begin{aligned} f(x + h) &= T_{f,x,n}(h) + \frac{h^{n+1}}{(n + 1)(h - \eta h)^n} \frac{f^{(n+1)}(x + \eta h)}{n!} (h - \eta h)^n \\ &= T_{f,x,n}(h) + \frac{f^{(n+1)}(x + \eta h)}{(n + 1)!} h^{n+1}, \end{aligned}$$

wie behauptet. ■

Definition: a) $T_{f,x,n}(h) = \sum_{k=0}^n \frac{h^k}{k!} f^{(k)}(x)$ heißt TAYLOR-Polynom n -ten

Grades von f um den Punkt x .

b) Der Fehlerterm $R_{n+1}(x, h) = \frac{h^{n+1}}{(n + 1)!} f^{(n+1)}(x + \eta h)$ heißt *Restglied* von LAGRANGE.

c) Ist f beliebig oft differenzierbar, bezeichnen wir die unendliche Reihe

$$T_{f,x}(h) = \sum_{k=0}^{\infty} \frac{1}{k!} f^{(k)}(x) h^k$$

als TAYLOR-Reihe von f um x .

d) Eine beliebig oft differenzierbare Funktion f heißt *analytisch* an der Stelle x , wenn es ein $\delta > 0$ gibt, so daß die TAYLOR-Reihe $T_{f,x}(h)$ für alle h mit $|h| < \delta$ gegen $f(x + h)$ konvergiert.



BROOK TAYLOR (1685–1731) war Sohn wohlhabender Eltern und wurde daher, bevor er 1703 an die Universität Cambridge ging, nur von privaten Hauslehrern ausgebildet. In Cambridge beschäftigte er sich hauptsächlich mit Mathematik, woran sich auch nach seinem Studienabschluß nichts änderte. Sein 1715 erschienenes Buch *Methodus incrementorum directa et inversa* enthält unter anderem TAYLOR-Polynome (die in Spezialfällen bereits LEIBNIZ, NEWTON und anderen bekannt waren), sowie die Methode der partiellen Integration. Weitere Bücher und Arbeiten beschäftigen sich unter anderem mit der Perspektive sowie mit Fragen aus der Physik.



JOSEPH-LOUIS LAGRANGE (1736–1813) wurde als GIUSEPPE LODOVICO LAGRANGIA in Turin geboren und studierte dort zunächst Latein. Erst eine alte Arbeit von HALLEY über algebraische Methoden in der Optik weckte sein Interesse an der Mathematik, woraus ein ausgedehnter Briefwechsel mit EULER entstand. In einem Brief vom 12. August 1755 berichtete er diesem unter anderem über seine Methode zur Berechnung von Maxima und Minima; 1756 wurde er, auf EULERS Vorschlag, Mitglied der Berliner Akademie; zehn Jahre später zog er nach Berlin und wurde dort EULERS Nachfolger als mathematischer Direktor der

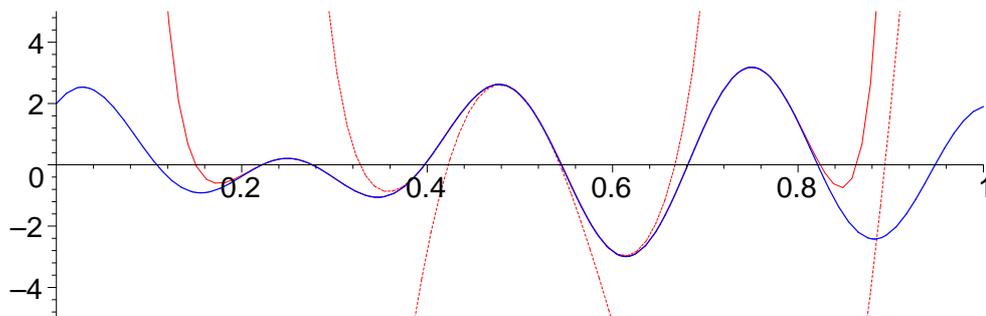
Akademie. 1787 wechselte er an die Pariser Académie des Sciences, wo er bis zu seinem Tod blieb und unter anderem an der Einführung des metrischen Systems beteiligt war. Seine Arbeiten umspannen weite Teile der Analysis, Algebra und Geometrie.

Natürlich kann ein TAYLOR-Polynom eines festen Grades eine Funktion, die nicht durch ein Polynom gegeben ist, nicht auf ganz \mathbb{R} gut approximieren: Ein Polynom n -ten Grades etwa hat höchstens n Nullstellen und wächst unbeschränkt, wenn x gegen $\pm\infty$ geht, während etwa eine Funktion wie der Sinus (mit dem wir uns im nächsten Paragraphen genauer beschäftigen werden) unendlich viele Nullstellen hat und für alle reellen Argumente Funktionswerte zwischen -1 und 1 hat.

Wir erwarten aber, daß bei einer hinreichend oft differenzierbaren Funktion die Übereinstimmung für große Grade für kleine h immer besser und für große h in einem immer größeren Bereich akzeptabel wird.

Als Beispiel ist hier der Graph einer Funktion abgebildet zusammen mit den Graphen der TAYLOR-Polynome zweiten Grades (gepunktet), achten

Grades (gestrichelt) und zwanzigsten Grades (durchgezogen). Dieses Bild entspricht durchaus der Vorstellung, die man sich im allgemeinen machen sollte.



Um TAYLOR-Polynome in einigen konkreten Fällen auszurechnen, betrachten wir als erstes Beispiel die Funktion $f(x) = e^x$ um den Punkt $x = 0$. Da $f'(x) = f(x) = e^x$ ist, sind auch alle höheren Ableitungen $f^{(k)}(x) = e^x$; somit ist

$$e^{x+h} = \sum_{k=0}^n \frac{h^k}{k!} e^x + \frac{h^{n+1}}{(n+1)!} e^{x+\eta h}$$

mit einer reellen Zahl $\eta \in (0, 1)$. Wenn wir durch e^x dividieren (oder $x = 0$ setzen), erhalten wir die Formel

$$e^h = \sum_{k=0}^n \frac{h^k}{k!} + \frac{h^{n+1}}{(n+1)!} e^{\eta h}.$$

Die Folge der Zahlen $h^n/n!$ ist eine Nullfolge, denn für eine natürliche Zahl $m \geq |h|$ ist für $n > m$

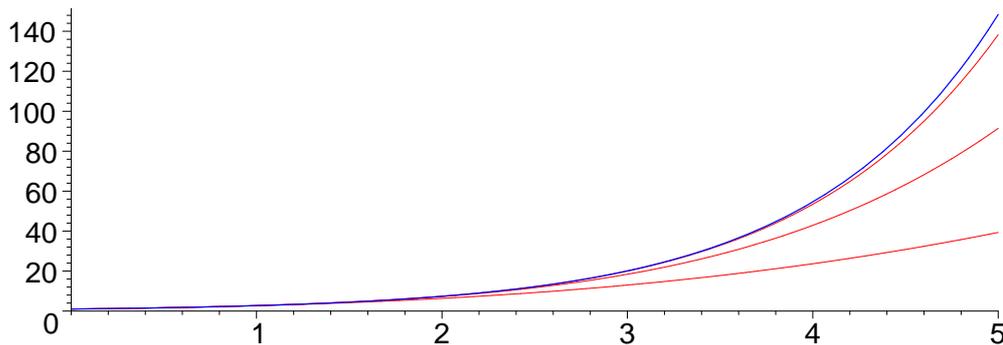
$$\frac{h^n}{n!} = \frac{h^m}{m!} \cdot \frac{h}{m+1} \cdot \frac{h}{m+2} \cdots \frac{h}{n} \leq \frac{h^m}{m!} \left(\frac{h}{m+1} \right)^{n-m},$$

und $|h/(m+1)| < 1$. Somit konvergiert die Reihe $\sum_{k=0}^{\infty} \frac{h^k}{k!}$ gegen e^h , wir bekommen, wenn wir die Variable h durch x ersetzen, die Formel

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!};$$

die Exponentialfunktion ist damit insbesondere analytisch. Mit der gerade bewiesenen Formel können wir die Werte von e^x mit beliebiger Genauigkeit ausrechnen – zumindest im Prinzip.

Da für $x > 0$ alle Summanden der TAYLOR-Reihe positiv sind, ist klar, daß im positiven Bereich die Werte der TAYLOR-Polynome stets kleiner sind als der Wert der Exponentialfunktion und daß TAYLOR-Polynome höheren Grades größere Werte liefern als die kleineren Grades. Die Abbildung zeigt dies für die Grade drei, fünf und acht; wie man sieht ist die Übereinstimmung für $0 \leq x \leq 5$ bei Grad acht schon so gut, daß sich die beiden Kurven kaum unterscheiden lassen:



Bevor wir übermütig werden, wollen wir nach dieser Formel mit zehnstelliger Genauigkeit den Wert von e^{-30} berechnen. Als Ergebnis erhalten wir $e^{-30} \approx 878229229,3$.

Das kann eigentlich kaum stimmen: Bekanntlich ist $e \approx 2,718281828$ größer als zwei, und $2^{30} = 1024^3 > 1000^3 = 10^9$, also sollte $e^{-30} < 10^{-9}$ sein. Der Grund für dieses katastrophale Ergebnis sind natürlich Rundungsfehler: Wir berechnen eine alternierende Summe mit immer kleiner werdenden Termen ausgehend vom größten Term. Die Rundungsfehler der ersten Additionen dominieren somit die immer kleiner werdenden hinteren Terme.

Für positive Werte von x konvergiert die Reihe sehr gut; das Problem mit den negativen Werten wird dadurch umgangen, daß man für $x > 0$ den Wert von e^{-x} berechnet als $1/e^x$.

Betrachten wir als nächstes Beispiel die Hyperbelfunktionen

$$\sinh x = \frac{e^x - e^{-x}}{2} \quad \text{und} \quad \cosh x = \frac{e^x + e^{-x}}{2} .$$

Offensichtlich ist die Ableitung des *Sinus hyperbolicus* der *Cosinus hyperbolicus* ist und umgekehrt; für $f(x) = \sinh x$ ist also

$$f^{(n)}(x) = \begin{cases} \sinh x & \text{für gerade } n \\ \cosh x & \text{für ungerade } n \end{cases} ;$$

speziell für $x = 0$ verschwinden alle geraden Ableitungen und die ungeraden haben den Wert eins. Das TAYLOR-Polynom vom Grad $2n$ ist daher

$$T_{\sinh,0,2n}(h) = \sum_{k=0}^{n-1} \frac{h^{2k+1}}{(2k+1)!},$$

und das Restglied ist

$$R_{2n+1}(0, h) = \frac{h^{2n+1}}{(2n+1)!} \cosh \eta h$$

mit einer reellen Zahl η zwischen 0 und 1. Die Zahl η hängt dabei natürlich von n ab, aber da der *Cosinus hyperbolicus* im Positiven monoton steigend und im Negativen monoton fallend ist, ist auf jeden Fall $\cosh(\eta h) \leq \cosh h$. Damit ist

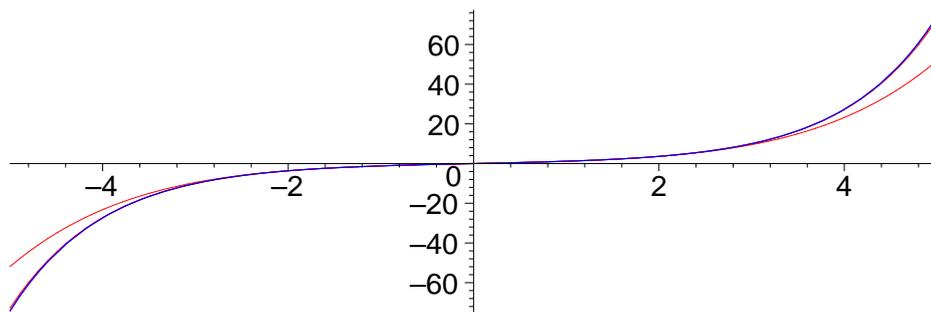
$$|R_{2n+1}(0, h)| = \left| \frac{h^{2n+1}}{(2n+1)!} \cosh(\eta h) \right| \leq \frac{|h|^{2n+1}}{(2n+1)!} \cosh h,$$

und rechts steht für festes h und wachsendes n eine Nullfolge. Daher ist nicht nur für jedes n

$$\sinh x = \sum_{k=0}^{n-1} \frac{x^{2k+1}}{(2k+1)!} + R_{2n+1}(0, x),$$

sondern auch im Limes $\sinh x = \sum_{k=0}^{\infty} \frac{x^{2k+1}}{(2k+1)!}$.

Das nachstehende Bild zeigt die Funktion zusammen mit ihren TAYLOR-Polynomen zweiten, vierten und sechsten Grades. Das erstere ist unsichtbar, weil es sich praktisch nicht von der x -Achse unterscheidet, das letzte, weil es sich praktisch nicht von der Kurve unterscheidet.



Im Sinne der obigen Definitionen ist also der *Sinus hyperbolicus* analytisch im Nullpunkt; tatsächlich überzeugt man sich leicht, daß er, wie die Exponentialfunktion, sogar analytisch auf ganz \mathbb{R} ist. Genauso zeigt man, daß auch der *Cosinus hyperbolicus* analytisch ist und

$$\cosh x = \sum_{k=0}^{\infty} \frac{x^{2k}}{(2k)!}.$$

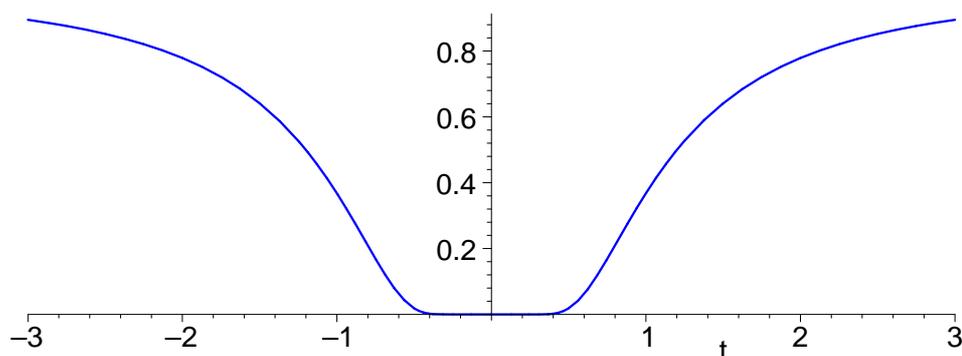
Eine Funktion kann höchstens dann analytisch in x sein, wenn sie dort beliebig oft differenzierbar ist – andernfalls läßt sich ihre TAYLOR-Reihe nicht einmal definieren. Analytizität ist aber eine noch stärkere Eigenschaft als beliebige Differenzierbarkeit, wie das folgende Beispiel zeigt: Wir setzen

$$f(x) = \begin{cases} e^{-1/x^2} & \text{für } x \neq 0 \\ 0 & \text{für } x = 0 \end{cases}.$$

Diese Funktion ist stetig, denn für $x \rightarrow 0$ geht $-1/x^2$ gegen $-\infty$, e^{-1/x^2} also gegen null. Auch mit der Differenzierbarkeit gibt es keine Probleme, denn für $x \neq 0$ ist $f'(x) = \frac{2}{x^3}e^{-1/x^2}$, und wie man leicht induktiv zeigen kann, haben auch alle weiteren Ableitungen die Form

$$f^{(n)}(x) = \text{rationale Funktion} \times e^{-1/x^2}.$$

Die rationale Funktion geht zwar gegen unendlich für $x \rightarrow 0$, aber da e^{-1/x^2} viel schneller gegen null geht als irgendeine rationale Funktion gegen unendlich, ist der Grenzwert des Produkts für $x \rightarrow 0$ gleich null. Also ist f auch im Nullpunkt beliebig oft differenzierbar, und alle Ableitungen verschwinden. Die TAYLOR-Reihe von f um den Nullpunkt ist daher die Nullfunktion, d.h. die Reihe konvergiert zwar überall, aber außerhalb des Nullpunkts nicht gegen e^{-1/x^2} .



Wie das Bild der Funktion f zeigt, ist sie in der Umgebung des Nullpunkts in der Tat sehr flach und unterscheidet sich kaum von der Null; danach steigt sie an in Richtung Eins, der sie sich dann –immer flacher werdend – für $x \rightarrow \pm\infty$ von unten her annähert.

Bislang sind wir von einer Funktion ausgegangen und haben versucht, diese durch eine Potenzreihe darzustellen. Manchmal ist auch das umgekehrte nützlich. Als Beispiel dazu betrachten wir ein Problem, das FIBONACCI in seinem 1202 erschienenen Buch *Liber abaci* präsentierte:

Ein Mann bringt ein Paar Karnickel auf einen Platz, der von allen Seiten durch eine Mauer umgeben ist. Wie viele Paare können von diesem Paar innerhalb eines Jahres produziert werden, wenn man annimmt, daß jedes Paar jeden Monat ein neues Paar liefert, das vom zweiten Monat nach seiner Geburt an produktiv ist?



LEONARDO PISANO (1170–1250) ist heute vor allem unter seinem Spitznamen FIBONACCI bekannt; gelegentlich nannte er sich auch BIGOLLO, auf Deutsch *Tunichtgut* oder *Reisender*. Er ging in Nordafrika zur Schule, kam aber 1202 zurück nach Pisa. Seine Bücher waren mit die ersten, die die indisch-arabischen Ziffern in Europa einführten. Er behandelt darin nicht nur Rechenaufgaben für Kaufleute, sondern auch zahlentheoretische Fragen, beispielsweise daß man die Quadratzahlen durch Aufaddieren der ungeraden Zahlen erhält. Auch betrachtet er Beispiele nichtlinearer Gleichungen, die er approximativ löst, und erinnert an viele in Vergessenheit geratene Ergebnisse der antiken Mathematik.

Wenn wir dies in Formeln übersetzen, sind also zunächst $F_0 = 0$ Paare von Karnickeln auf dem Platz; im ersten Monat wird eines hingbracht, so daß es nun $F_1 = 1$ sind. Im zweiten Monat kommen keine neuen Karnickel dazu, also sind es weiterhin $F_2 = 1$ Paare. Zu Beginn des dritten Monat allerdings ist das vorhandene Paar zwei Monate alt und produziert somit ein neues Paar, wir haben also $F_3 = 2$ Paare. Das gleiche geschieht zu Beginn des vierten Monats, ab dem fünften Monat ist auch das zu Beginn des dritten Monats geborene Paar aktiv. Allgemein können

wir sagen, daß zu Beginn des k -ten Monats alle Paare, die spätestens zu Beginn des $(k - 1)$ -ten Monats geboren wurden, Nachkommen liefern; da FIBONACCI nichts über sterbende Karnickel sagt, ist außerdem die gesamte Population des Vormonats noch vorhanden, wir haben also F_{k-1} alte und F_{k-2} neue Paare. Die Anzahl von Karnickeln im k -ten Monat ist damit durch eine rekursiv definierte Folge $(F_k)_{k \in \mathbb{N}_0}$ gegeben mit

$$F_0 = 0, \quad F_1 = 1 \quad \text{und} \quad F_k = F_{k-1} + F_{k-2} \quad \text{für } k \geq 2.$$

Damit läßt sich F_k zumindest grundsätzlich ausrechnen; für große Werte von k wäre es allerdings besser, wenn wir eine einfache Formel hätten, die uns ein festes F_k liefert, ohne daß wir auch dessen sämtliche Vorgänger berechnen müssen.

Dazu betrachten wir, zunächst noch ohne jede Frage nach der Konvergenz, die Potenzreihe

$$R(x) = \sum_{k=0}^{\infty} F_k x^k.$$

Die Bedingung $F_k = F_{k-1} + F_{k-2}$ können wir durch diese Reihe ausdrücken, indem wir sie mit x beziehungsweise x^2 multiplizieren:

$$\begin{aligned} xR(x) &= \sum_{k=0}^{\infty} F_k x^{k+1} = \sum_{k=1}^{\infty} F_{k-1} x^k \\ \text{und} \\ x^2 R(x) &= \sum_{k=0}^{\infty} F_k x^{k+2} = \sum_{k=2}^{\infty} F_{k-2} x^k. \end{aligned}$$

Somit ist

$$\begin{aligned} xR(x) + x^2 R(x) &= \sum_{k=1}^{\infty} F_{k-1} x^k + \sum_{k=2}^{\infty} F_{k-2} x^k \\ &= \sum_{k=2}^{\infty} (F_{k-1} + F_{k-2}) x^k = \sum_{k=2}^{\infty} F_k x^k, \end{aligned}$$

denn $F_0 = 0$. Aus diesem Grund und weil $F_1 = 1$ ist, können wir $R(x)$ schreiben als

$$R(x) = x + \sum_{k=2}^{\infty} F_k x^k, \quad \text{also ist} \quad R(x) = x + xR(x) + x^2 R(x).$$

Fassen wir alle Terme mit $R(x)$ zusammen, gibt uns dies die Beziehung

$$(1 - x - x^2)R(x) = x \quad \text{oder} \quad R(x) = \frac{x}{1 - x - x^2}.$$

Da wir uns bislang nicht um Konvergenz gekümmert haben, waren alle diese Umformungen nur spekulativ; wir können jetzt aber versuchen, die erhaltene rationale Funktion durch eine Potenzreihe darzustellen und zu untersuchen, ob deren Koeffizienten die Definitionsgleichungen der F_k erfüllen.

Dazu versuchen wir, wie bereits mehrfach bei der Untersuchung von Reihen mit Summanden einer ähnlichen Form, die Funktion als Summe zweier Terme mit linearen Nennern zu schreiben. Wenn wir diese Nenner in der Form $1 - q$ schreiben können, liefert uns die Summenformel der geometrischen Reihe eine Potenzreihenentwicklung.

Bestimmen wir also zunächst die Nullstellen des Nenners:

$$1 - x - x^2 = -\left(x + \frac{1}{2}\right)^2 + \frac{5}{4} = 0 \quad \text{für} \quad x = x_{1/2} = -\frac{1}{2} \pm \frac{\sqrt{5}}{2}.$$

Da $x_1 x_2 = -1$ ist, können wir dies auch schreiben als

$$1 - x - x^2 = \frac{x - x_1}{x_1} \cdot \frac{x - x_2}{x_2} = \left(1 - \frac{x}{x_1}\right) \left(1 - \frac{x}{x_2}\right).$$

Dabei ist

$$\frac{1}{x_1} = \frac{2}{-1 + \sqrt{5}} = \frac{2(1 + \sqrt{5})}{5 - 1} = \frac{1 + \sqrt{5}}{2} \quad \text{und} \quad \frac{1}{x_2} = \frac{1 - \sqrt{5}}{2}.$$

Diese beiden Zahlen werden in der Mathematik traditionellerweise mit

$$\phi = \frac{1}{2} + \frac{\sqrt{5}}{2} \approx 1,618033988 \quad \text{und} \quad \bar{\phi} = \frac{1}{2} - \frac{\sqrt{5}}{2} \approx -0,618033988$$

bezeichnet; ϕ ist die Zahl des seit der Antike vielfach in der Kunst verwendeten *goldenen Schnitts*. Sie erfüllen die quadratische Gleichung

$$1 - \frac{1}{x} - \frac{1}{x^2} = 0 \quad \text{oder} \quad x^2 - x - 1 = 0.$$

Dividieren wir die daraus resultierende Gleichung $\phi^2 = \phi + 1$ durch ϕ , erhalten wir

$$\phi = \frac{\phi + 1}{\phi};$$

stehen also zwei Strecken a und b im Verhältnis ϕ , so ist

$$\frac{a}{b} = \phi = \frac{\phi + 1}{\phi} = \frac{\frac{a+b}{b}}{\frac{a}{b}} = \frac{a+b}{a};$$

die größere Strecke a steht daher im gleichen Verhältnis zur kleineren Strecke b wie die Summe $a + b$ zur größeren Strecke a .

Für $\bar{\phi}$ gilt natürlich formal dasselbe; da $\bar{\phi}$ negativ ist, war das allerdings für Künstler nicht weiter interessant.

Nach der dritten binomischen Formel ist $\phi\bar{\phi} = -\frac{1}{4} = -1$; deshalb hat das quadratische Polynom $(1 - \phi x)(1 - \bar{\phi} x)$ ebenfalls die Nullstellen ϕ und $\bar{\phi}$; da es auch den konstanten Koeffizienten eins hat, ist also

$$1 - x - x^2 = (1 - \phi x)(1 - \bar{\phi} x),$$

und wir haben den Nenner unserer rationalen Funktion zerlegt in zwei Linearfaktoren der Form $1 - q$.

Wir wollen die Funktion darstellen in der Form

$$\frac{x}{1 - x - x^2} = \frac{a}{1 - \phi x} + \frac{b}{1 - \bar{\phi} x} = \frac{(a+b) - (a\bar{\phi} + b\phi)x}{1 - x - x^2}.$$

Dazu muß $a + b = 0$ und $a\bar{\phi} + b\phi = -1$ sein, also $b = -a$ und damit $a(\bar{\phi} - \phi) = -a\sqrt{5} = -1$, d.h. $a = 1/\sqrt{5}$. Somit ist

$$\frac{x}{1 - x - x^2} = \frac{1/\sqrt{5}}{1 - \phi x} - \frac{1/\sqrt{5}}{1 - \bar{\phi} x}.$$

Auf diese beiden Summanden können wir nun die Summenformel für geometrische Reihen anwenden und erhalten für $|x| < 1/\phi$ die Darstellung

$$\frac{x}{1 - x - x^2} = \frac{1}{\sqrt{5}} \left(\sum_{k=0}^{\infty} \phi^k x^k - \sum_{k=0}^{\infty} \bar{\phi}^k x^k \right) = \sum_{k=0}^{\infty} \frac{\phi^k - \bar{\phi}^k}{\sqrt{5}} x^k.$$

Um zu sehen, daß die Koeffizienten dieser Potenzreihe wirklich die FIBONACCI-Zahlen sind, müssen wir zeigen, daß sie deren Definitionsgleichungen erfüllen: Für $k = 0$ und $k = 1$ ist

$$\frac{\phi^0 - \bar{\phi}^0}{\sqrt{5}} = \frac{1 - 1}{\sqrt{5}} = 0 \quad \text{und} \quad \frac{\phi - \bar{\phi}}{\sqrt{5}} = \frac{\left(\frac{1}{2} + \frac{\sqrt{5}}{2}\right) - \left(\frac{1}{2} - \frac{\sqrt{5}}{2}\right)}{\sqrt{5}} = 1,$$

wie gewünscht. Für $k \geq 2$ ist zunächst

$$\phi^k = \phi^{k-2} \cdot \phi^2 = \phi^{k-2} \cdot (\phi + 1) = \phi^{k-1} + \phi^{k-2}$$

und entsprechend für $\bar{\phi}$, da beides Lösungen der quadratischen Gleichung $x^2 = x + 1$ sind. Setzen wir beides ein in die Formel für die Koeffizienten der Potenzreihe, erhalten wir die gewünschte Beziehung

$$\begin{aligned} \frac{\phi^k - \bar{\phi}^k}{\sqrt{5}} &= \frac{\phi^{k-1} + \phi^{k-2} - \bar{\phi}^{k-1} - \bar{\phi}^{k-2}}{\sqrt{5}} \\ &= \frac{\phi^{k-1} - \bar{\phi}^{k-1}}{\sqrt{5}} + \frac{\phi^{k-2} - \bar{\phi}^{k-2}}{\sqrt{5}}. \end{aligned}$$

Damit erfüllen die Koeffizienten dieser Potenzreihe die definierenden Gleichungen der FIBONACCI-Zahlen, sind also mit diesen identisch. Für alle $k \in \mathbb{N}_0$ ist also

$$F_k = \frac{\phi^k - \bar{\phi}^k}{\sqrt{5}} = \frac{\phi^k}{\sqrt{5}} - \frac{\bar{\phi}^k}{\sqrt{5}}.$$

Der zweite Summand hat für alle $k \geq 0$ einen Betrag kleiner als $\frac{1}{2}$, denn $\bar{\phi} \approx -0,618$ und $\sqrt{5} \approx 2,236$. Da alle FIBONACCI-Zahlen ganz sind, können wir F_k deshalb auch einfacher ausrechnen als die nächste ganze Zahl zu $\phi^k / \sqrt{5}$. Insbesondere steigt die Folge der FIBONACCI-Zahlen also exponentiell.

Entsprechend kann man auch bei anderen linearen Rekursionen der Form

$$a_k = c_1 a_{k-1} + c_2 a_{k-2} + \dots + c_r a_{k-r}$$

vorgehen. Das Prinzip sei hier nur kurz skizziert, da man mit Methoden der Linearen Algebra mit weniger Aufwand ein besseres Ergebnis beweisen kann.

Für eine Rekursion, bei der a_k von seinen r Vorgängern abhängt, müssen natürlich die ersten r Folgenglieder a_0, \dots, a_{r-1} explizit gegeben sein; erst der Rest läßt sich dann über die Rekursionsformel berechnen. Betrachten wir auch hier wieder die Potenzreihe

$$R(x) = \sum_{k=0}^{\infty} a_k x^k,$$

so erhalten wir nun die Gleichung

$$R(x) = P(x) + c_1 x R(x) + c_2 x^2 R(x) + \cdots + c_r x^r R(x),$$

wobei $P(x)$ ein Polynom vom Grad höchstens $r - 1$ ist, das von den r Anfangstermen abhängt. Auflösen nach $R(x)$ führt auf die rationale Funktion

$$R(x) = \frac{P(x)}{1 - c_1 x - c_2 x^2 - \cdots - c_r x^r}.$$

Falls das Nennerpolynom r verschiedene (reelle oder komplexe) Nullstellen x_1, \dots, x_r hat, können es schreiben als

$$1 - c_1 x - c_2 x^2 - \cdots - c_r x^r = \prod_{i=1}^r \left(1 - \frac{x}{x_i}\right)$$

und wir können $R(x)$ zerlegen als

$$R(x) = \sum_{i=1}^r \frac{e_i}{1 - \frac{x}{x_i}}.$$

Hier steht x_i im Nenner; um es in den Zähler zu bekommen, dividieren wir einfach das Polynom, dessen Nullstellen die x_i sind, durch x^r :

$$\frac{1}{x^r} - c_1 \cdot \frac{1}{x^{r-1}} - c_2 \cdot \frac{1}{x^{r-2}} - \cdots - c_{r-1} \frac{1}{x} - c_r$$

verschwindet genau dann, wenn wir für x eines der x_i einsetzen. Somit sind die Zahlen $1/x_i$ die Nullstellen des Polynoms

$$y^r - c_1 y^{r-1} - c_2 y^{r-2} - \cdots - c_r.$$

Bezeichnen wir diese mit y_1, \dots, y_r , ist also

$$R(x) = \sum_{i=1}^r \frac{e_i}{1 - y_i x}.$$

Ist $|y_i x| < 1$ für alle i , können wir dies nach der Summenformel für die geometrische Reihe ausdrücken als

$$\sum_{i=1}^r e_i \sum_{k=0}^{\infty} (y_i x)^k = \sum_{k=0}^{\infty} \sum_{i=1}^r e_i y_i^k x^k.$$

Wie oben zeigt man, daß die Koeffizienten b_k dieser Potenzreihe die Rekursionformel $b_k = c_1 b_{k-1} + c_2 b_{k-2} + \cdots + c_r b_{k-r}$ erfüllen; damit

folgt, daß sich die a_k als Linearkombinationen der Zahlen y_i^k schreiben lassen. Da die ersten r der a_k explizit gegeben sind, führen die ersten r der Gleichungen

$$y_1^k e_1 + y_2^k e_2 + \cdots + y_r^k e_r = a_k$$

auf ein lineares Gleichungssystem aus r Gleichungen für die r Unbekannten e_i , aus dem sich diese bestimmen lassen; wir müssen die rationale Funktion $R(x)$ also nicht explizit zerlegen.

Beim Beispiel der FIBONACCI-Zahlen ist $r = 2$ und wir haben die beiden Nullstellen ϕ und $\bar{\phi}$. Deren Kehrwerte sind $-\bar{\phi}$ und $-\phi$, also ist F_k eine Linearkombination der Form $F_k = e_1 \phi^k + e_2 \bar{\phi}^k$. Da $F_0 = 0$ und $F_1 = 1$ ist, führt das auf die beiden Gleichungen

$$0 = e_1 + e_2 \quad \text{und} \quad 1 = e_1 \phi + e_2 \bar{\phi} = \frac{e_1 + e_2}{2} + \frac{(e_1 - e_2)\sqrt{5}}{2}.$$

Somit ist

$$e_2 = -e_1 \quad \text{und} \quad e_1 = \frac{1}{\sqrt{5}}, \quad \text{also} \quad F_k = \frac{\phi^k - \bar{\phi}^k}{\sqrt{5}},$$

wie beim ersten Ansatz.

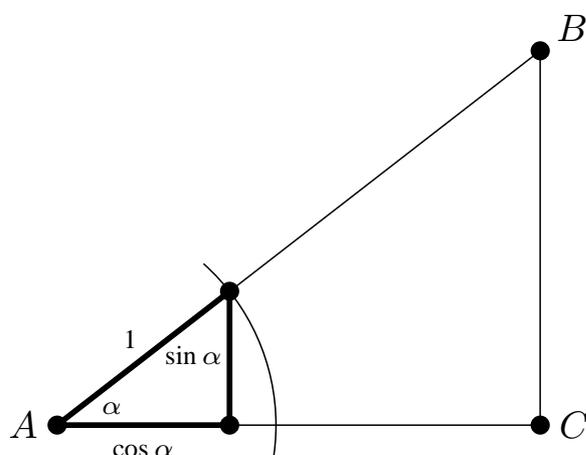
§4: Trigonometrische Funktionen

Unter den aus der Schule bekannten Funktionen gibt es noch eine große Klasse von Funktionen, mit denen wir uns hier noch überhaupt nicht beschäftigt haben, die trigonometrischen Funktionen. In diesem Paragraphen sollen sie zunächst so definiert werden, wie das üblicherweise in der Schule passiert; danach wollen wir ihre analytischen Eigenschaften studieren.

a) Klassische geometrische Definition

Definition: Das bei C rechtwinklige Dreieck mit Ecken A, B, C habe bei A den Winkel α . Dann ist

$$\sin \alpha \stackrel{\text{def}}{=} \frac{\text{Gegenkathete}}{\text{Hypotenuse}} = \frac{|\overline{BC}|}{|\overline{AB}|} \quad \text{und} \quad \cos \alpha \stackrel{\text{def}}{=} \frac{\text{Ankathete}}{\text{Hypotenuse}} = \frac{|\overline{AC}|}{|\overline{AB}|}.$$



Da zwei rechtwinklige Dreiecke mit einem Winkel α bis auf eine eventuell notwendige Spiegelung ähnlich zueinander sind, hängt diese Definition in der Tat nur vom Winkel α ab und nicht auch noch von dem gewählten Dreieck. In der Tat kann man Sinus und Kosinus auch direkt sehen, indem man ein rechtwinkliges Dreieck mit Hypotenuse der Länge eins betrachtet; dann ist $\sin \alpha$ gleich der Länge der Gegenkathete und $\cos \alpha$ gleich der der Ankathete. Insbesondere zeigt dann der Satz des PYTHAGORAS, daß

$$\sin^2 \alpha + \cos^2 \alpha = 1$$

ist für jeden Winkel α , wobei wir hier wie auch im folgenden die bei trigonometrischen Funktionen allgemein übliche Konvention

$$\sin^n \alpha \stackrel{\text{def}}{=} (\sin \alpha)^n \quad \text{und} \quad \cos^n \alpha \stackrel{\text{def}}{=} (\cos \alpha)^n$$

verwenden.

Wenn wir im Dreieck $\triangle ABC$ die Rolle der beiden Ecken A und B vertauschen, haben wir dort den Winkel $90^\circ - \alpha$, und bezüglich dieses Winkels sind die Rollen von Ankathete und Gegenkathete vertauscht. Daher ist

$$\sin(90^\circ - \alpha) = \cos \alpha \quad \text{und} \quad \cos(90^\circ - \alpha) = \sin \alpha .$$

Dies erklärt übrigens auch den Namen *Kosinus*: Es handelt sich um den *Sinus* des Komplementärwinkels. (Das lateinische Wort *Sinus* bedeutet *Bucht* oder *Meerbusen*: es kam in die Mathematik als falsche lateinische Übersetzung des arabischen Worts für *Halbsehne*.)

Zumindest für spezielle Winkel lassen sich die so definierten Winkel-funktionen mit elementargeometrischen Methoden exakt ausrechnen: Der Fall $\alpha = 0^\circ$ etwas entspricht einem rechtwinkligen „Dreieck“, das zu einem waagrechten Strich mit $B = C$ degeneriert ist; da dann $\overline{AB} = \overline{AC}$ ist, während die Strecke \overline{BC} die Länge null hat, ist somit

$$\sin 0^\circ = \cos 90^\circ = 0 \quad \text{und} \quad \sin 90^\circ = \cos 0^\circ = 1.$$

Für $\alpha = 45^\circ$ wird das rechtwinklige Dreieck gleichschenkelig, so daß Si-nus und Kosinus von 45° übereinstimmen müssen. Da nach PYTHAGO-RAS die Quadratsumme der beiden gleich eins ist, folgt

$$\sin 45^\circ = \cos 45^\circ = \sqrt{\frac{1}{2}} = \frac{\sqrt{2}}{2}.$$

Für $\alpha = 30^\circ$ bzw. $\alpha = 60^\circ$ ist das rechtwinklige Dreieck gleich der oberen bzw. linken Hälfte eines gleichseitigen Dreiecks, die Gegenkathete bzw. Ankathete ist also halb so lang wie die Hypotenuse und damit ist

$$\sin 30^\circ = \cos 60^\circ = \frac{1}{2}$$

und nach PYTHAGORAS

$$\cos 30^\circ = \sin 60^\circ = \sqrt{1 - \left(\frac{1}{2}\right)^2} = \frac{\sqrt{3}}{2}.$$

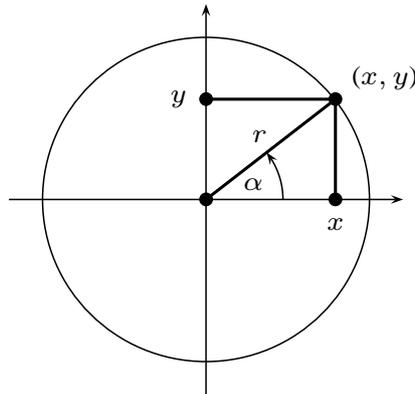
Mit etwas Mühe könnten wir auf diese Weise noch viele andere Funk-tionswerte berechnen; in der Tat erstellte bereits um 140 vor Christus der griechische Mathematiker HIPPARCHOS VON RHODOS (180–125) mit solchen Methoden eine (leider nicht erhaltene) zwölfbändige Tafel von Sinuswerten.

b) Allgemeine Definition

Mit der bisherigen Definition können wir nur Winkel zwischen 0° und 90° betrachten; ein Sinus von 150° läßt sich so nicht erklären, denn es gibt schließlich kein rechtwinkliges Dreieck mit einem Winkel von 150° .

Betrachten wir aber auf dem Einheitskreis, d.h. dem Kreis mit Radius eins um den Nullpunkt O , jenen Punkt $P = (x, y)$, für den der Radius \overline{OP} mit der x -Achse den Winkel φ einschließt, so haben dieser für

$0^\circ \leq \varphi \leq 90^\circ$ ein rechtwinkliges Dreieck mit Ecken O , P und $(x, 0)$; nach obiger Definition ist daher $x = \cos \varphi$ und $y = \sin \varphi$, denn die Hypotenuse hat als Radius die Länge eins, die Ankathete hat Länge x und die Gegenkathete Länge y .



Diese Interpretation führt sofort zu einer Definition von Sinus und Kosinus für beliebige Winkel:

Definition: Für jeden Winkel $\varphi \in \mathbb{R}$ ist der Sinus die y -Koordinate und der Kosinus die x -Koordinate jenes Punktes P auf der Einheitskreislinie, für den der zugehörige Radius mit der x -Achse den Winkel φ bildet.

Da wir nach 360° wieder am Ausgangspunkt landen, folgt daraus insbesondere, daß die so definierten Funktionen periodisch sind mit einer Periode von 360° .

Die so definierten Winkelfunktionen stimmen allerdings immer noch nicht ganz mit den üblicherweise in der Analysis benutzten Funktionen gleichen Namens überein: In der Analysis werden Winkel nicht im Gradmaß angegeben, sondern im sogenannten *Bogenmaß*. Das Bogenmaß eines Winkels ist derjenige Bogen auf dem Umfangs des Einheitskreises, der zum Kreissektor mit diesem Winkel gehört.

Aus Symmetriegründen muß das Bogenmaß eines Winkels proportional sein zum Gradmaß; da der volle Kreisumfang einerseits einem Winkel von 360° entspricht, andererseits aber einem Bogenmaß von 2π , ist also

$$\text{Bogenmaß} = \frac{2\pi}{360^\circ} \times \text{Gradmaß} = \frac{\pi}{180^\circ} \times \text{Gradmaß}.$$

Insbesondere sind Sinus und Kosinus in der neuen Interpretation nun periodisch mit der Periode 2π :

$$\sin(x + 2k\pi) = \sin x \quad \text{und} \quad \cos(x + 2k\pi) = \cos x \quad \text{für alle } k \in \mathbb{Z}.$$

Einige wichtige spezielle Werte für Winkel im Bogenmaß sind

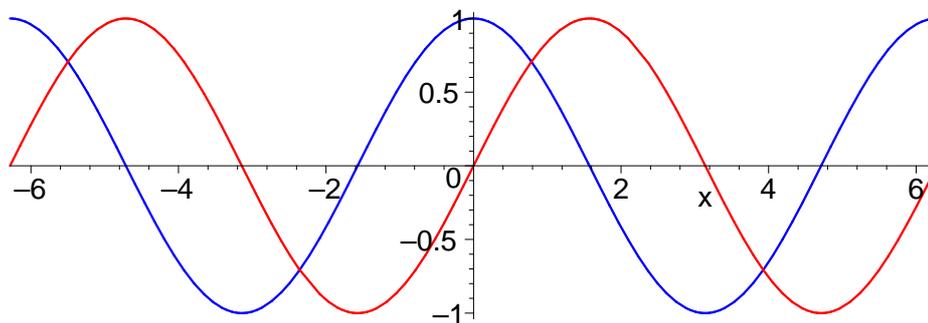
0°	30°	45°	60°	90°	120°	135°	150°	180°
0	$\frac{\pi}{6}$	$\frac{\pi}{4}$	$\frac{\pi}{3}$	$\frac{\pi}{2}$	$\frac{2\pi}{3}$	$\frac{3\pi}{4}$	$\frac{5\pi}{6}$	π

Wenn wir im folgenden von $\sin x$ oder $\cos x$ sprechen, soll x **immer als Winkel im Bogenmaß** aufgefaßt werden. Die so definierten Funktionen sind sowohl für die mathematische Theorie als auch für Anwendungen wie der Modellierung saisonaler Trends erheblich nützlicher als die entsprechenden Funktionen des Winkels im Gradmaß. Vor allem für Studenten, die gerne mit dem Taschenrechner arbeiten, ist es wichtig, immer an die Unterscheidung zwischen Winkelmaß und Gradmaß zu denken: Die meisten Taschenrechner arbeiten standardmäßig mit dem Gradmaß, was in Klausuren immer wieder zu unsinnigen Ergebnissen führt. (Taschenrechner bieten meist auch noch sogenannte *Neugrad* an, die so definiert sind, daß ein rechter Winkel hundert Neugrad hat. Praktische Anwendungen dieser Neugrad sind mir nicht bekannt.) Bei Computern gibt es weniger Probleme, da die meisten Compiler- und Interpreterbibliotheken trigonometrische Funktionen nur für Argumente im Bogenmaß zur Verfügung stellen. Mit der neuen Definition haben wir die folgenden speziellen Werte von Sinus und Kosinus:

x	0	$\frac{\pi}{6}$	$\frac{\pi}{4}$	$\frac{\pi}{3}$	$\frac{\pi}{2}$
$\sin x$	0	$\frac{1}{2}$	$\frac{1}{2}\sqrt{2}$	$\frac{1}{2}\sqrt{3}$	1
$\cos x$	1	$\frac{1}{2}\sqrt{3}$	$\frac{1}{2}\sqrt{2}$	$\frac{1}{2}$	0 .

Als Eselsbrücke kann man sich dies auch merken in der Form

x	0	$\frac{\pi}{6}$	$\frac{\pi}{4}$	$\frac{\pi}{3}$	$\frac{\pi}{2}$
$\sin x$	$\frac{1}{2}\sqrt{0}$	$\frac{1}{2}\sqrt{1}$	$\frac{1}{2}\sqrt{2}$	$\frac{1}{2}\sqrt{3}$	$\frac{1}{2}\sqrt{4}$
$\cos x$	$\frac{1}{2}\sqrt{4}$	$\frac{1}{2}\sqrt{3}$	$\frac{1}{2}\sqrt{2}$	$\frac{1}{2}\sqrt{1}$	$\frac{1}{2}\sqrt{0}$

Sinus und Kosinus zwischen -2π und 2π

c) Die Ableitungen der trigonometrischen Funktionen

Da $\cos x = \sin\left(\frac{\pi}{2} - x\right)$ ist, genügt es, den Sinus abzuleiten. Wir berechnen seine Ableitung als Grenzwert des Differenzenquotienten

$$\frac{\sin(x+h) - \sin x}{h}.$$

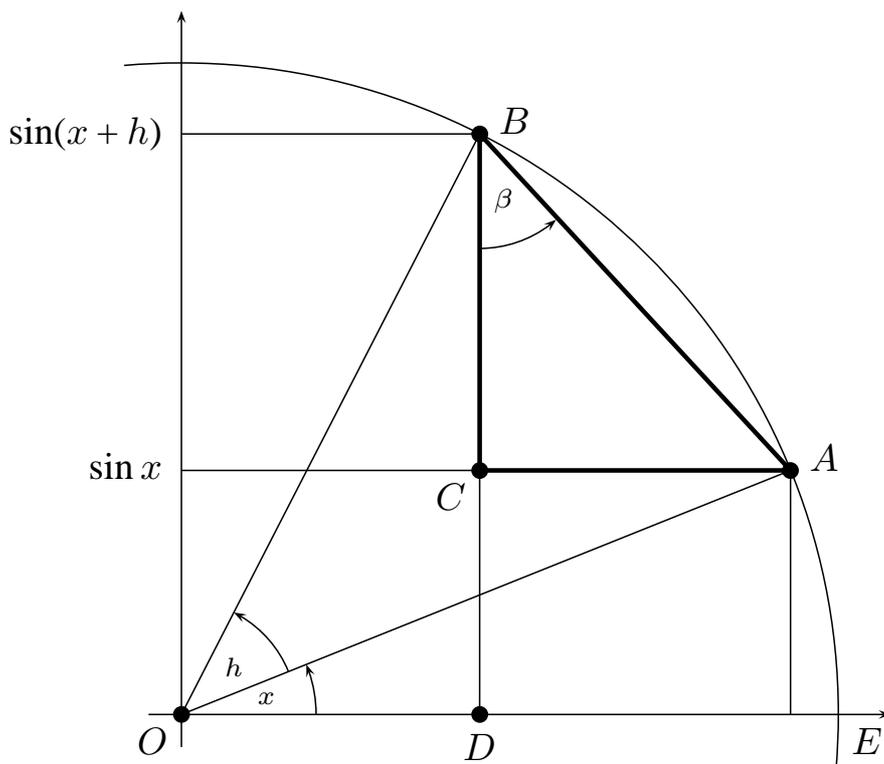
Betrachten wir die Situation geometrisch: Wir tragen x als Winkel zur reellen Achse ab dem Nullpunkt auf; in der Abbildung ist x also gleich der Länge des Bogens von $E = (1, 0)$ nach A . Entsprechend ist $x+h$ gleich der Länge des Bogens von E nach B , und h selbst ist die des Bogens von A nach B .

Nach Definition ist $\sin x$ die Ordinate von A und $\sin(x+h)$ die von B ; obiger Zähler $\sin(x+h) - \sin x$ ist also die Länge der Strecke \overline{CB} . Der Nenner h ist die Länge des Bogens von A nach B . Da wir Bogenlängen über immer feiner unterteilte Streckenzüge definieren können, unterscheidet sich dieser Bogen für immer kleiner werdendes h beliebig wenig von der Strecke \overline{AB} , d.h.

$$\lim_{h \rightarrow 0} \frac{\sin(x+h) - \sin x}{h} = \lim_{h \rightarrow 0} \frac{|\overline{CB}|}{|\overline{AB}|}.$$

Da $\triangle ABC$ rechtwinklig ist, ist dieses Streckenverhältnis gerade der Kosinus des Winkels β bei B , und wir müssen diesen berechnen.

Dazu betrachten wir zunächst das Dreieck $\triangle OAB$. Da A und B auf dem Einheitskreis liegen, ist es gleichschenkelig und hat somit bei A und B denselben Winkel. Da die Winkelsumme im Dreieck gleich π und der Winkel bei O gleich h ist, errechnet sich dieser zu $\frac{\pi-h}{2}$.



Als nächstes betrachten wir das rechtwinklige Dreieck $\triangle OBD$. Da es bei O den Winkel $x + h$ hat, muß der Winkel bei B gleich

$$\pi - \frac{\pi}{2} - (x + h) = \frac{\pi}{2} - (x + h)$$

sein, und β als Differenz dieser beiden Winkel ergibt sich zu

$$\beta = \frac{\pi - h}{2} - \frac{\pi}{2} + x + h = x + \frac{h}{2}.$$

Somit ist die Ableitung von $\sin x$ an der Stelle x gleich

$$\lim_{h \rightarrow 0} \cos \left(x + \frac{h}{2} \right) = \cos x,$$

die Ableitung des Sinus ist also der Kosinus. Die Ableitung des Kosinus ist die von $\sin \left(\frac{\pi}{2} - x \right)$, also nach der Kettenregel

$$-\cos \left(\frac{\pi}{2} - x \right) = -\sin x.$$

d) Die Differentialgleichung $f''(x) = -\omega^2 f(x)$

Werden Sinus oder Kosinus zweimal abgeleitet, so wird je einmal ein Sinus und ein Kosinus abgeleitet; im letzteren Fall kehrt sich das Vorzei-

chen um, im ersteren bleibt es gleich. Daher ist die zweite Ableitung von Sinus bzw. Kosinus gerade die negativ genommene jeweilige Funktion.

Allgemein genügen für jedes $\omega \in \mathbb{R}$ sowohl $f(x) = \cos \omega x$ als auch $f(x) = \sin \omega x$ der Gleichung $f''(x) = -\omega^2 f(x)$.

Diese Gleichung sollte, vor allem wenn man x durch die Zeit t ersetzt, aus der Physik bekannt sein: Sie beschreibt die Schwingungen eines Gewichts an einer Feder, deren Rückstellkraft sich aus dem HOOKEschen Gesetz ergibt, die Schwingungen eines Pendels bei kleiner Auslenkung, den Strom in einem ungedämpften elektrischen Schwingkreis usw.; sie tritt also in vielen Zusammenhängen auf, in denen wir es mit periodischen Vorgängen zu tun haben. Da Zykeln auch im Wirtschaftsleben eine große Rolle spielen, sollten wir sie etwas genauer betrachten.

Da $\sin \omega x$ und $\cos \omega x$ Lösungen sind, ist auch jede Linearkombination

$$f(x) = a \cos \omega x + b \sin \omega x \quad \text{mit} \quad a, b \in \mathbb{R}$$

eine Lösung. Das Hauptergebnis dieses Paragraphen besagt, daß dies im Fall $\omega \neq 0$ bereits *alle* Lösungen sind: (Was gilt für $\omega = 0$?)

Satz: Die in einem offenen Intervall (r, s) mit $r < 0 < s$ definierte reelle Funktion f genüge der Differentialgleichung $f''(x) = -\omega^2 f(x)$ mit einem nichtverschwindenden ω . Dann gibt es Konstanten $a, b \in \mathbb{R}$, so daß $f(x) = a \cos \omega x + b \sin \omega x$ ist.

Beweis: Wir multiplizieren die Differentialgleichung mit $2f'(x)$ und erhalten

$$2f'(x)f''(x) = -2\omega^2 f(x)f'(x).$$

Nach der Produkt- oder Kettenregel können wir die linke Seite auch auffassen als die Ableitung von $f'(x)^2$ und die rechte als Ableitung von $-\omega^2 f(x)^2$; die Ableitung der Funktion $f'(x)^2 + \omega^2 f(x)^2$ verschwindet somit überall. Es gibt daher eine Konstante $c \in \mathbb{R}$, so daß

$$f'(x)^2 + \omega^2 f(x)^2 = c$$

ist für alle $x \in (r, s)$. Wir unterscheiden zwei Fälle:

Erster Fall: $c = 0$. Da Quadrate reeller Zahlen nie negativ sein können, muß dann für jedes x gelten $f(x) = f'(x) = 0$, und mit $a = b = 0$ läßt

sich die Nullfunktion in der Tat als Linearkombination von $\cos \omega x$ und $\sin \omega x$ schreiben, womit der Satz für diesen Fall bewiesen wäre.

Zweiter Fall: $c \neq 0$. Dieser Fall soll auf den ersten zurückgeführt werden.

Falls wir annehmen, der Satz sei richtig, gibt es $a, b \in \mathbb{R}$ so daß gilt $f(x) = a \cos \omega x + b \sin \omega x$. Alsdann ist $f(0) = a$ und $f'(0) = b\omega$, also

$$a = f(0) \quad \text{und} \quad b = \frac{f'(0)}{\omega}.$$

Für eine beliebige Lösung f der Differentialgleichung $f''(x) = -\omega f(x)$ definieren wir reelle Zahlen a, b über diese Gleichungen und betrachten die Funktion

$$h(x) = f(x) - f(0) \cos \omega x - \frac{f'(0)}{\omega} \sin \omega x.$$

Auch h ist, wie man leicht nachrechnet, eine Lösung der Differentialgleichung, und nach Konstruktion ist $h(0) = h'(0) = 0$. Daher verschwindet für h die oben definierte Konstante c , die man ja über jeden beliebigen x -Wert berechnen kann, und nach dem, was wir im ersten Fall gezeigt haben, ist h gleich der Nullfunktion. Damit ist aber

$$f(x) = f(0) \cos \omega x + \frac{f'(0)}{\omega} \sin \omega x,$$

wie behauptet. ■

Die Voraussetzung, daß 0 im Intervall (r, s) enthalten sein soll, ist nicht wirklich notwendig; sie macht nur den Beweis bequemer und übersichtlicher. Statt mit den Werten von $f(0)$ und $f'(0)$ könnte man auch mit $f(u)$ und $f'(u)$ für ein beliebiges $u \in (r, s)$ argumentieren, müßte dann allerdings zur Bestimmung von a und b noch ein lineares Gleichungssystem lösen.

Für die Modellierung eines Systems ist der Satz meist in der folgenden Form etwas nützlicher:

Korollar: Die in einem offenen Intervall (r, s) definierte reelle Funktion f genüge der Differentialgleichung $f''(x) = -\omega^2 f(x)$ mit einem nichtverschwindenden ω , und für einen Wert $x_0 \in (r, s)$ sei $f(x_0) = a$ und $f'(x_0) = b$. Dann ist

$$f(x) = a \cos \omega(x - x_0) + \frac{b}{\omega} \sin \omega(x - x_0).$$

Beweis: Für $x_0 = 0$ und $r < 0 < s$ haben wir das bereits oben im Beweis des Satzes gezeigt, und die Verschiebung des Arguments um x_0 ändert natürlich nichts. ■

e) Die Taylorreihen der trigonometrischen Funktionen

Nachdem wir aus Abschnitt c) die ersten Ableitungen von Sinus und Kosinus kennen, bereiten uns auch die höheren Ableitungen keine Schwierigkeiten mehr: Die Ableitung des Sinus ist der Kosinus, dessen Ableitung wiederum $-\text{Sinus}$, die Ableitung davon ist $-\text{Kosinus}$, und die Ableitung davon wieder der Sinus, womit alles wieder von vorne losgeht. Formal aufgeschrieben ist also die n -te Ableitung der Sinusfunktion gleich

$$\sin^{(n)} x = \begin{cases} \sin x & \text{falls } n \equiv 0 \pmod{4} \\ \cos x & \text{falls } n \equiv 1 \pmod{4} \\ -\sin x & \text{falls } n \equiv 2 \pmod{4} \\ -\cos x & \text{falls } n \equiv 3 \pmod{4}. \end{cases}$$

Dabei soll $a \equiv b \pmod{m}$ bedeuten, daß $a - b$ ohne Rest durch m teilbar ist oder, anders ausgedrückt, daß a und b bei der Division durch m den gleichen Rest haben.

Entsprechend ist

$$\cos^{(n)} x = \begin{cases} \cos x & \text{falls } n \equiv 0 \pmod{4} \\ -\sin x & \text{falls } n \equiv 1 \pmod{4} \\ -\cos x & \text{falls } n \equiv 2 \pmod{4} \\ \sin x & \text{falls } n \equiv 3 \pmod{4}. \end{cases}$$

Da $\sin 0 = 0$ und $\cos 0 = 1$ ist, sind damit auch die Werte aller höherer Ableitungen an der Stelle $x = 0$ bekannt: Für gerade n verschwindet $\sin^{(n)}(0)$, für ungerade $n = 2k + 1$ ist $\sin^{(n)}(0) = (-1)^k$; entsprechend verschwindet $\cos^{(n)}(0)$ für ungerade n , und für gerade $n = 2k$ ist $\cos^{(n)}(0) = (-1)^k$. Die TAYLOR-Polynome sind somit

$$\begin{aligned} T_{\sin,0,2n+1}(h) &= T_{\sin,0,2n+2}(h) = \sum_{k=0}^n (-1)^k \frac{h^{2k+1}}{(2k+1)!} \\ &= h - \frac{h^3}{6} + \frac{h^5}{120} - \dots + (-1)^n \frac{h^{2n+1}}{(2n+1)!} \end{aligned}$$

und

$$\begin{aligned} T_{\cos,0,2n}(h) &= T_{\cos,0,2n+1}(h) = \sum_{k=0}^n (-1)^k \frac{h^{2k}}{(2k)!} \\ &= 1 - \frac{h^2}{2} + \frac{h^4}{24} - \cdots + (-1)^n \frac{h^{2n}}{(2n)!}. \end{aligned}$$

Noch nicht ganz klar ist allerdings, ob die als Limes dieser Polynome erhaltenen TAYLOR-Reihen die Funktionen Sinus und Kosinus tatsächlich darstellen; bei einigen Funktionen wie etwa $f(x) = e^{-1/x^2}$ kann es ja durchaus vorkommen, daß die TAYLOR-Reihe gegen eine andere Funktion als die Ausgangsfunktion konvergiert; im Beispiel war dies die Nullfunktion.

Hier ist das glücklicherweise nicht der Fall: Nach dem Satz über die TAYLOR-Entwicklung ist

$$\sin(0+h) = T_{\sin,0,2n+2}(h) + R_{2n+3}(h)$$

mit

$$R_{2n+3}(h) = (-1)^{n+1} \frac{\sin^{(2n+3)}(\eta h)}{(2n+3)!} h^{2n+3} \quad \text{für ein } \eta \in (0, 1).$$

Da der Sinus nur Werte vom Betrag höchstens eins annimmt, können wir das Restglied abschätzen durch

$$|R_{2n+3}(h)| \leq \frac{|h|^{2n+3}}{(2n+3)!}.$$

Wie wir schon bei der TAYLOR-Reihe der Exponentialfunktion gesehen haben, geht $h^k/k!$ für alle $h \in \mathbb{R}$ gegen Null für $k \rightarrow \infty$, also geht auch hier das Restglied gegen Null für $n \rightarrow \infty$, d.h.

$$\sin x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!} = x - \frac{x^3}{6} + \frac{x^5}{120} - \cdots,$$

wobei ich hier anstelle von h den Buchstaben x verwendet habe. Genauso folgt auch die Formel

$$\cos x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!} = 1 - \frac{x^2}{2} + \frac{x^4}{24} - \cdots.$$

f) Die Eulerschen Formeln und ihre Anwendung

Die TAYLOR-Reihen für Sinus und Kosinus erinnern von ihrer Struktur her an die der Exponentialfunktion: Der einzige Unterschied besteht darin, daß jeweils nur „die Hälfte“ der Terme auftritt und daß das Vorzeichen alterniert.

Um den Zusammenhang zwischen den drei Reihen wirklich zu verstehen, müssen wir sie im Komplexen betrachten. Wir überlegen uns zunächst, daß die Reihen

$$\sum_{k=0}^{\infty} \frac{z^k}{k!}, \quad \sum_{k=0}^{\infty} (-1)^k \frac{z^{2k+1}}{(2k+1)!} \quad \text{und} \quad \sum_{k=0}^{\infty} (-1)^k \frac{z^{2k}}{(2k)!}$$

für alle $z \in \mathbb{C}$ konvergieren: Wie wir aus §6 des zweiten Kapitels wissen, konvergiert eine Reihe mit komplexen Summanden insbesondere dann, wenn sie absolut konvergent ist, wenn also die Reihe mit den Absolutbeträgen der Summanden konvergiert. In unseren drei Fällen sind dies die (reellen) Potenzreihen

$$\sum_{k=0}^{\infty} \frac{|z|^k}{k!}, \quad \sum_{k=0}^{\infty} \frac{|z|^{2k+1}}{(2k+1)!} \quad \text{und} \quad \sum_{k=0}^{\infty} \frac{|z|^{2k}}{(2k)!},$$

und das sind TAYLOR-Reihen, von denen wir bereits gezeigt haben, daß sie gegen $e^{|z|}$, $\sinh |z|$ beziehungsweise $\cosh |z|$ konvergieren. Somit konvergieren auch die obigen Reihen.

Definition: Für eine komplexe Zahl $z \in \mathbb{C}$ ist $e^z = \sum_{k=0}^{\infty} \frac{z^k}{k!}$,

$$\sin z = \sum_{k=0}^{\infty} (-1)^k \frac{z^{2k+1}}{(2k+1)!} \quad \text{und} \quad \cos z = \sum_{k=0}^{\infty} (-1)^k \frac{z^{2k}}{(2k)!}.$$

Für reelle Werte von z stehen hier gerade die TAYLOR-Reihen dieser Funktionen, und wir wissen, daß diese gegen die Funktionswerte konvergieren. Die obige Definition führt also für reelle Zahlen zum gleichen Ergebnis wie die bisherigen Definitionen.

Wenn wir die Exponentialfunktion für eine rein imaginäre Zahl $z = ix$

berechnen wollen, erhalten wir

$$\begin{aligned} e^{ix} &= \sum_{k=0}^{\infty} \frac{(ix)^k}{k!} = \sum_{\ell=0}^{\infty} (-1)^\ell \frac{x^{2\ell}}{(2\ell)!} + i \sum_{\ell=0}^{\infty} (-1)^\ell \frac{x^{2\ell+1}}{(2\ell+1)!} \\ &= \cos x + i \sin x. \end{aligned}$$

Entsprechend ist $e^{-ix} = \cos x - i \sin x$, also

$$\cos x = \Re e^{ix} = \frac{e^{ix} + e^{-ix}}{2} \quad \text{und} \quad \sin x = \Im e^{ix} = \frac{e^{ix} - e^{-ix}}{2i}.$$

Die gerade bewiesenen Formeln werden als EULERSche Formeln bezeichnet; mit ihnen lassen sich die trigonometrischen Funktionen auf die erheblich handlichere Exponentialfunktion zurückzuführen.

Die wichtigste Eigenschaft der reellen Exponentialfunktion ist die Funktionalgleichung $e^{x+y} = e^x \cdot e^y$; wir sollten uns deshalb überlegen, ob diese auch im Komplexen gilt. Wir müssen also entscheiden, ob für beliebige komplexe Zahlen z, w gilt

$$\sum_{k=0}^{\infty} \frac{z^k}{k!} \cdot \sum_{\ell=0}^{\infty} \frac{w^\ell}{\ell!} = \sum_{m=0}^{\infty} \frac{(z+w)^m}{m!}.$$

Wenn wir einfach naiv darauf losrechnen, ist

$$\begin{aligned} \sum_{k=0}^{\infty} \frac{z^k}{k!} \cdot \sum_{\ell=0}^{\infty} \frac{w^\ell}{\ell!} &= \sum_{k=0}^{\infty} \sum_{\ell=0}^{\infty} \frac{z^k w^\ell}{k! \ell!} = \sum_{m=0}^{\infty} \left(\sum_{k=0}^m \frac{z^k w^{m-k}}{k! (m-k)!} \right) \\ &= \sum_{m=0}^{\infty} \left(\sum_{k=0}^m \binom{m}{k} \frac{z^k w^{m-k}}{m!} \right) = \sum_{m=0}^{\infty} \frac{(z+w)^m}{m!}, \end{aligned}$$

wie gewünscht. Wie wir am Beispiel der Reihe $\sum_{k=1}^{\infty} (-1)^k$ gesehen haben, können solche Umformungen allerdings bei unendlichen Reihen zu widersprüchlichen und daher unsinnigen Ergebnissen führen; der obigen Rechnung ist also nicht zu trauen. Um sie zu rechtfertigen, müssen wir beide Seiten als Grenzwerte von Folgen auffassen und zeigen, daß beide Folgen gegen denselben Wert konvergieren, daß es also zu jedem $\varepsilon > 0$ ein $N \in \mathbb{N}$ gibt, so daß

$$\left| \sum_{k=0}^n \frac{z^k}{k!} \cdot \sum_{\ell=0}^n \frac{w^\ell}{\ell!} - \sum_{m=0}^n \frac{(z+w)^m}{m!} \right| < \varepsilon \quad (*)$$

ist für alle $n \geq N$. Da die Funktionalgleichung im Reellen gilt, wissen wir, daß dies für alle *reellen* Werte von z und w der Fall ist. Nach obiger Rechnung ist der abzuschätzende Betrag gleich

$$\left| \sum_{m=n+1}^{2n} \sum_{k=m-n}^n \binom{m}{k} \frac{z^k w^{m-k}}{m!} \right|,$$

und nach der Dreiecksungleichung ist dies höchstens gleich dem Betrag von

$$\sum_{m=n+1}^{2n} \sum_{k=m-n}^n \binom{m}{k} \frac{|z|^k |w|^{m-k}}{m!},$$

also dem Betrag der Differenz

$$\sum_{k=0}^n \frac{|z|^k}{k!} \cdot \sum_{\ell=0}^n \frac{|w|^\ell}{\ell!} - \sum_{m=0}^n \frac{(|z| + |w|)^m}{m!}.$$

Da $e^{|z|}e^{|w|} = e^{|z|+|w|}$ ist, wissen wir, daß es zu jedem $\varepsilon > 0$ ein $N \in \mathbb{N}$ gibt, so daß *diese* Summe für $n \geq N$ kleiner als ε ist. Damit gilt aber nach der Dreiecksungleichung für alle $n \geq N$ auch die Ungleichung (*), und dies beweist die Funktionalgleichung $e^{z+w} = e^z \cdot e^w$ für beliebige *komplexe* Zahlen z und w .

Als erste Anwendung betrachten wir die eher unhandlichen Additionsformeln für die trigonometrischen Funktionen: Für die Exponentialfunktion ist natürlich $e^{i(x+y)} = e^{ix}e^{iy}$. Nach den EULERSchen Formeln ist die linke Seite gleich $\cos(x+y) + i \sin(x+y)$ und die rechte ist

$$\begin{aligned} & (\cos x + i \sin x)(\cos y + i \sin y) \\ &= \cos x \cos y - \sin x \sin y + i(\sin x \cos y + \cos x \sin y). \end{aligned}$$

Daher ist

$$\cos(x+y) = \cos x \cos y - \sin x \sin y$$

und

$$\sin(x+y) = \sin x \cos y + \cos x \sin y.$$

Entsprechend lassen sich auch $\cos nx$ und $\sin nx$ berechnen.

Mit der Multiplikationsformel können wir auch den Wert von e^z für jede komplexe Zahl $z = x+iy$ durch bekannte reelle Funktionen ausdrücken: Nach der gerade bewiesenen Funktionalgleichung und den EULERSchen

Formeln ist $e^z = e^x(\cos y + i \sin y)$. In der Klammer steht eine komplexe Zahl vom Betrag $\sqrt{\cos^2 y + \sin^2 y} = 1$; daher ist der Betrag von e^z gleich e^x . Da die reelle Exponentialfunktion die reellen Zahlen bijektiv abbildet auf die positiven reellen Zahlen, kann der Betrag von e^z somit auch für komplexe z nie verschwinden, die Exponentialfunktion nimmt also auch im Komplexen nie den Wert Null an.

Ist $w = u + iv$ eine komplexe Zahl vom Betrag eins, so liegt der Punkt $(u, v) \in \mathbb{R}^2$ auf der Einheitskreislinie; somit gibt es einen Winkel y derart, daß $u = \cos y$ und $v = \sin y$ ist, also $w = e^{iy}$. Hat w einen beliebigen Betrag $r > 0$, so gibt es ein $x \in \mathbb{R}$ mit $e^x = r$ und, da w/r den Betrag eins hat, ein $y \in \mathbb{R}$ mit $e^{iy} = w/r$. Für die komplexe Zahl $z = x + iy$ ist daher $e^z = e^{x+iy} = e^x \cdot e^{iy} = w$. Damit haben wir gezeigt, daß es zu jeder komplexen Zahl $w \in \mathbb{C} \setminus \{0\}$ eine komplexe Zahl z gibt mit $e^z = w$; die komplexe Exponentialfunktion bildet also \mathbb{C} surjektiv ab auf $\mathbb{C} \setminus \{0\}$. Im Gegensatz zur reellen Exponentialfunktion ist allerdings die komplexe Exponentialfunktion nicht mehr injektiv: Da Sinus und Kosinus periodisch sind mit Periode 2π , ist $e^z = e^{z+2k\pi i}$ für alle $k \in \mathbb{Z}$. Insbesondere ist $e^{2\pi i} = 1$. Erwähnenswert ist auch die Gleichung

$$e^{i\pi} = \cos \pi + i \sin \pi = -1,$$

die mit e , π , dem Minuszeichen und der Eins gleich vier wesentliche Symbole der Mathematik in Beziehung zueinander setzt. Sie ist übrigens auch der Ausgangspunkt für den 1882 veröffentlichten Beweis von CARL LOUIS FERDINAND VON LINDENMANN (1852–1939), daß π nicht nur eine irrationale Zahl ist, sondern sogar *transzendent*; das bedeutet, daß π im Gegensatz etwa zu $\sqrt{2}$ keine Nullstelle eines Polynoms mit rationalen Koeffizienten ist. Für e hatte dies CHARLES HERMITE (1822–1901) bereits 1873 direkt bewiesen; für π benutzte LINDENMANN die Gleichung $e^{\pi i} = -1$. Aus dem Beweis von LINDENMANN folgt insbesondere, daß das klassische Problem der Quadratur des Kreises mit Zirkel und Lineal unlösbar ist.

Die EULERSchen Formeln erlauben auch eine neue Art der Darstellung komplexen Zahlen: Jede komplexe Zahl $w \neq 0$ läßt sich darstellen in der Form $w = e^z = e^x \cdot e^{iy}$ mit einer komplexen Zahl $z = x + iy$, wobei $r = e^x$ der Betrag von w ist. Da $|0| = 0$ ist, läßt sich daher *jede* komplexe Zahl z darstellen in der Form $z = r e^{iy}$ mit einer nichtnegativen reellen

Zahl r und einer reellen Zahl φ . Die Zahl r ist einfach der Betrag von z ; die Zahl φ bezeichnen wir als *Argument* von z , in Zeichen $\varphi = \arg z$. Es ist, wie alle Winkel, nur bis auf ganzzahlige Vielfache von 2π eindeutig bestimmt (und für $z = 0$ sogar völlig unbestimmt). Die Darstellung $z = re^{i\varphi}$ von z heißt *Polarkoordinatendarstellung* von z .

In dieser Polarkoordinatendarstellung wird die Multiplikation und Division komplexer Zahlen erheblich einfacher als in der klassischen kartesischen Darstellung: $re^{i\varphi} \cdot se^{i\psi} = (rs)e^{i(\varphi+\psi)}$, d.h. bei der Multiplikation werden die Beträge miteinander multipliziert und die Argumente addiert. Dabei zeigt sich auch, daß die Unbestimmtheit des Arguments durchaus ihren Sinn hat: Die Gleichung $(-i) \cdot (-i) = -1$ wird in Polarkoordinatendarstellung zu $e^{3\pi i/2} \cdot e^{3\pi i/2} = e^{3\pi i} = e^{\pi i} = -1$.

Auch im Reellen ist die Polarkoordinatendarstellung gelegentlich nützlich: Statt durch die kartesischen Koordinaten (x, y) kann man die Lage eines Punktes in \mathbb{R}^2 auch festlegen durch seinen Abstand r vom Nullpunkt sowie den Winkel φ zwischen ihrer Verbindungsstrecke mit dem Nullpunkt und der x -Achse. Die kartesischen Koordinaten lassen sich aus r und φ leicht berechnen über die Formeln $x = r \cos \varphi$ und $y = r \sin \varphi$; umgekehrt ist $r = \sqrt{x^2 + y^2}$, und φ muß so berechnet werden, daß $\cos \varphi = x/r$ und $\sin \varphi = y/r$ ist.

Als letzte Anwendung trigonometrischer Funktion möchte ich noch eine Funktion angeben, die zwar differenzierbar, zumindest im Nullpunkt aber nicht stetig differenzierbar ist: Wir setzen

$$f(x) = \begin{cases} x^2 \cos \frac{1}{x} & \text{für } x \neq 0 \\ 0 & \text{für } x = 0 \end{cases}.$$

Für $x \neq 0$ haben wir nach den üblichen Rechenregeln die Ableitung

$$f'(x) = 2x \cos \frac{1}{x} - x^2 \sin \frac{1}{x} \cdot \frac{-1}{x^2} = 2x \cos \frac{1}{x} + \sin \frac{1}{x},$$

und $f'(0) = \lim_{h \rightarrow 0} \frac{h^2 \cos \frac{1}{h} - 0}{h} = \lim_{h \rightarrow 0} h \cos \frac{1}{h} = 0$; die Ableitung existiert also für alle $x \in \mathbb{R}$. Sie ist aber im Nullpunkt nicht stetig, denn $\lim_{x \rightarrow 0} \sin \frac{1}{x}$ existiert nicht einmal.

§5: Zusammenfassung

Wir haben in diesem Kapitel differenzierbare Funktionen kennengelernt als Funktionen, die sich „im Kleinen“ wie lineare Funktionen verhalten. Dazu gehören die meisten „gängigen“ mathematischen Funktionen, insbesondere alles, was aus Grundrechenarten und Funktionen wie der Exponentialfunktion, dem Logarithmus und den trigonometrischen Funktionen zusammengesetzt ist, solange man diese Funktionen nur anwendet, wo sie definiert sind und natürlich auch nie durch die Null dividiert.

Für solche Funktionen f haben wir die *Ableitung* f' definiert als die Steigung derjenigen Geraden, durch die sich die Funktion am besten approximieren läßt, also der Tangenten. f heißt *glatt*, wenn f' eine stetige Funktion ist.

Für grundlegende Funktionen konnten wir die Ableitungen ausrechnen; die wichtigsten sind in der folgenden Tabelle zusammengefaßt:

$$\begin{array}{l} f(x) = \quad x^n, n \in \mathbb{R} \quad e^x \quad \log x \quad \sin x \quad \cos x \\ f'(x) = \quad nx^{n-1} \quad e^x \quad \frac{1}{x} \quad \cos x \quad -\sin x \end{array}$$

Wir lernten auch Rechenregeln kennen, über die sich Ableitungen zusammengesetzter Funktionen berechnen lassen:

$$\begin{aligned} (f \pm g)'(x) &= f'(x) \pm g'(x), & (fg)'(x) &= f'(x)g(x) + f(x)g'(x) \\ \text{und} \quad \left(\frac{f}{g}\right)'(x) &= \frac{f'(x)g(x) - f(x)g'(x)}{g(x)^2}. \end{aligned}$$

Außerdem gilt für die Hintereinanderausführung von Funktionen die *Kettenregel* $(f \circ g)'(x) = f'(g(x))g'(x)$.

Falls eine differenzierbare Funktion $f: (a, b) \rightarrow (c, d)$ zwei (möglicherweise unendliche) Intervalle bijektiv aufeinander abbildet, gibt es eine *Umkehrfunktion* $g: (c, d) \rightarrow (a, b)$; sie ist differenzierbar in allen Punkten $y = f(x)$ mit $f'(x) \neq 0$, und $g'(y) = 1/f'(x)$.

Für viele differenzierbare Funktionen $f(x)$ ist auch die Ableitung $f'(x)$ differenzierbar; deren Ableitung bezeichnen wir als die *zweite Ableitung* $f''(x)$, und entsprechend definieren wir, sofern sie existieren, auch höhere Ableitungen.

Für eine mindestens n -fach differenzierbare Funktion können wir die Funktionswerte in der Umgebung eines Punktes x approximieren durch das TAYLOR-Polynom $T_{f,x,n}(h) = \sum_{k=0}^n \frac{f^{(k)}(x)}{k!} h^k$; falls f sogar $(n+1)$ -fach differenzierbar ist, gibt es ein $\eta \in (0, 1)$, so daß der Fehler dieser Approximation gegeben ist durch

$$f(x+h) - T_{f,x,n}(h) = R_{n+1,f}(h) = \frac{f^{(n+1)}(x+\eta h)}{(n+1)!}.$$

Für eine beliebig oft differenzierbare Funktion, bei der $R_{n+1,f}(h)$ für $n \rightarrow \infty$ zumindest für alle h mit einem Betrag kleiner einer geeigneten Zahl δ gegen Null geht, ist

$$f(x+h) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x)}{k!} h^k \quad \text{für alle } h \text{ mit } |h| < \delta.$$

Solche Funktionen heißen *analytisch*. Wichtige analytische Funktionen sind

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}, \quad \cos x = \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k}}{(2k)!} \quad \text{und} \quad \sin x = \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k+1}}{(2k+1)!};$$

über diese Reihen lassen sich die genannten Funktionen auch für beliebige *komplexe* Argumente definieren. Polynomfunktionen sind ebenfalls analytisch.

Die trigonometrischen Funktionen Sinus und Kosinus sind klassisch definiert über Winkel in eine rechtwinkligen Dreieck: Ist α dort einer der Winkel, ist

$$\sin \alpha = \frac{\text{Gegenkathete}}{\text{Hypotenuse}} \quad \text{und} \quad \cos \alpha = \frac{\text{Ankathete}}{\text{Hypotenuse}}.$$

Über die Koordinaten von Punkten auf dem Einheitskreis haben wir diese Funktionen erweitert zu Funktionen auf ganz \mathbb{R} und gesehen, daß sie via den Umweg über komplexe Zahlen mit den EULERSchen Formeln in Beziehung zur Exponentialfunktion gebracht werden können:

$$e^{ix} = \cos x + i \sin x, \quad \sin x = \frac{e^{ix} - e^{-ix}}{2i} \quad \text{und} \quad \cos x = \frac{e^{ix} + e^{-ix}}{2}.$$

Mit Hilfe dieser Formeln lassen sich viele klassische Beziehungen zwischen trigonometrischen Funktionen einfach herleiten beziehungsweise beweisen.

§6: Summary

In this chapter, we learned about differentiable functions: They are those functions that behave locally like linear functions. Most of the "usual" function are differentiable, in particular all those which are composed of basic arithmetic operations (except division by zero), exponential functions, logarithms (for positive arguments) and trigonometric functions.

The derivative f' of such a function f is defined as the slope of the line best approximation the function, that is, the slope of the tangent. f is called *smooth*, if f' is a continuous function.

For many basic functions we could determine their derivatives: the most important ones can be found in the following table:

$$\begin{array}{l} f(x) = \quad x^n, n \in \mathbb{R} \quad e^x \quad \log x \quad \sin x \quad \cos x \\ f'(x) = \quad nx^{n-1} \quad e^x \quad \frac{1}{x} \quad \cos x \quad -\sin x \end{array}$$

For more complex functions, the derivatives can be computing using the following rules:

$$\begin{aligned} (f \pm g)'(x) &= f'(x) \pm g'(x), & (fg)'(x) &= f'(x)g(x) + f(x)g'(x) \\ \text{and} \quad \left(\frac{f}{g}\right)'(x) &= \frac{f'(x)g(x) - f(x)g'(x)}{g(x)^2}. \end{aligned}$$

For compositions we have the *chain rule* $(f \circ g)'(x) = f'(g(x))g'(x)$.

For a differentiable function $f: (a, b) \rightarrow (c, d)$ mapping one (possibly infinite) interval bijectively to another, there exists an inverse function $g: (c, d) \rightarrow (a, b)$; it is differentiable in all points $y = f(x)$ where $f'(x) \neq 0$, and $g'(y) = 1/f'(x)$.

For many differentiable functions $f(x)$ also their derivative is differentiable; its derivative is called the *second* derivative of f and denoted $f''(x)$. Similarly, we can define higher derivatives, provided they exist.

If all derivatives up to the n^{th} exist, we can form the TAYLOR-polynomial

$T_{f,x,n}(h) = \sum_{k=0}^n \frac{f^{(k)}(x)}{k!} h^k$ as an approximation to f around x ; if also the $(n + 1)^{\text{st}}$ derivative exists, there exists an $\eta \in (0, 1)$ such that the error of this approximation is given by

$$f(x + h) - T_{f,x,n}(h) = R_{n+1,f}(h) = \frac{f^{(n+1)}(x + \eta h)}{(n + 1)!}.$$

If *all* derivatives of f exist around x , and if $\lim_{n \rightarrow \infty} R_{(n+1),f}(h) = 0$ for all h with $|h|$ smaller than some $\delta > 0$,

$$f(x + h) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x)}{k!} h^k \quad \text{for all } h \text{ such that } |h| < \delta.$$

Such functions are called *analytic*. Important analytic functions include

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}, \quad \cos x = \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k}}{(2k)!} \quad \text{and} \quad \sin x = \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k+1}}{(2k + 1)!};$$

using these series representations, we can define those functions also for complex arguments. Of course, polynomials are analytic, too.

The sine and cosine of an angle α are traditionally defined via right (or right-angled) triangles with angle α as

$$\sin \alpha = \frac{\text{opposite leg}}{\text{hypotenuse}} \quad \text{and} \quad \cos \alpha = \frac{\text{adjacent leg}}{\text{hypotenuse}}.$$

Using the coordinates of points on the unit circle, we could define those functions for arbitrary real arguments, and we found that via complex numbers they are related to the exponential function: By EULER's formulae we have

$$e^{ix} = \cos x + i \sin x, \quad \sin x = \frac{e^{ix} - e^{-ix}}{2i} \quad \text{and} \quad \cos x = \frac{e^{ix} + e^{-ix}}{2}.$$

Using these formulae, many relations between trigonometric functions can easily be derived or proven.